

THE FACTORIAL NATURE OF ATTAINMENT IN  
ELEMENTARY SCIENCE

By D. G. LEWIS

*(Department of Education and School of Education, University of Manchester)*

SUMMARY. Objective tests of attainment in physics, chemistry, and biology were administered to over three-hundred grammar-school pupils, together with a battery of verbal, number and spatial tests. A factorial analysis yielded the following results:

(1) The tests of physics, chemistry, and biology formed a group factor over and above a general factor.

(2) Verbal, spatial and numerical group factors were also found.

(3) Attainment in science, as measured by the objective tests, is dependent upon (a) the general factor, and (b) the scientific group factor.

(4) In addition, attainment in physics (though not in chemistry and biology) depends to some extent on the spatial factor.

(5) Again, for girls, attainment in biology depends to some extent on the verbal factor.

(6) Attainment in all three branches of school science is independent of the number factor. Numerical ability (as measured by speeded tests of the basic arithmetical processes) has no relation to attainment in school science.

(7) The factorial pattern underlying scientific attainment is very similar for boys and girls.

## I.—INTRODUCTION.

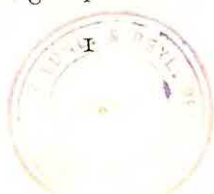
THIS investigation is concerned with an analysis of attainment in elementary science in terms of cognitive factors defined by psychological tests. Most of the previous studies in the field have been confined to analyses of attainment in a single science subject at about G.C.E. Ordinary Level or a little below. Thus, we have the studies of attainment in physics by Berridge (1947) and Jog (1955), in chemistry by Pawley (1937), in physics and chemistry by Chary (1948), in biology by James (1951) and Lamb (1954), and in botany by Ormiston (1939). The results taken as a whole seem to indicate that, in addition to the factor of general intellectual ability (*g*) and a general personality factor (such as the 'X' factor of Alexander's well-known investigation, 1935) verbal, spatial and numerical group factors are operating.

Analyses of attainment in science, mathematics and other school subjects (e.g., Wrigley, 1956; Lewis, 1961) show that science, together with mathematics, is distinguished from arts subjects in the first bipolar factor, while a subsequent factor separates the science subjects from mathematics.

It was thought that an analysis not confined to attainment factors and based on measures of all three major branches of school science was needed, and that this would have some bearing on the organization of science teaching in schools.

## II.—DESIGN OF EXPERIMENT.

Measures of attainment in physics, chemistry and biology were factor-analysed together with scores on psychological tests clearly defining verbal, spatial and numerical factors. It was thought that a suitable selection of tests would yield a general factor having at any rate a rough approximation to the general cognitive factor, *g*. Burt's group-factor method of analysis (1950) was



considered the most suitable as it retains a general factor and allows for possible overlapping among the group factors. The general design of the experiment is shown in Table 1. It was anticipated that a general factor together with verbal, spatial and numerical group factors would clearly emerge from the analysis. The questions marks in Table 1 would then correspond to the questions to be answered from the experiment, namely (i) is attainment in physics, chemistry and biology linked together by a group factor over and above a general factor? and (ii) to what extent does attainment in science, or in its separate branches, depend upon (a) verbal ability, (b) spatial ability, and (c) numerical ability?

TABLE 1

PLAN OF DESIGN OF EXPERIMENT.

TESTS	FACTORS				
	General	Verbal	Spatial	Numerical	Scientific
Verbal Tests .....	×	×			
	×	×			
	×	×			
Spatial Tests .....	×		×		
	×		×		
	×		×		
Number Tests .....	×			×	
	×			×	
	×			×	
Physics .....	×	?	?	?	?
Chemistry .....	×	?	?	?	?
Biology .....	×	?	?	?	?

### III.—DESCRIPTION OF TESTS.

Objective tests of attainment in physics, chemistry and biology were used. These tests were constructed by Mr. N. R. J. Simpson, a research student in the Department of Education in the Queen's University of Belfast, working in collaboration with the writer. The tests, which were based on the syllabus of the Northern Ireland Junior Certificate examination in experimental science, are primarily factual in nature, though some attempt was made to probe understanding of basic principles. Each test has two sections, a section of multiple-choice questions and a section of completion-type questions.\* All the multiple-choice questions required a choice from six possible answers. As the choice was so wide, it was considered that no correction for guessing need be made in the scoring. No penalization was made for incorrect spelling in the answers to the completion-type questions. If an answer was recognizably correct it was given full credit.

Simpson found test-retest reliabilities of .819 for physics, .860 for chemistry and .851 for biology. These results were obtained from a group of thirty-nine

\* Examples of items from the science tests are given in a previous paper (this *Journal*, XXXI, 111, 241-247). Mr. Simpson's permission to use the tests for the present study is gratefully acknowledged.



15-year-old boys with an interval of two weeks between the two test administrations. Simpson also found the correlation between the scores of 132 boys on the three tests combined (i.e., physics plus chemistry plus biology) and their scores in the experimental science examination (Junior Certificate) to be .752.

The verbal tests were all used by Thurstone in his well-known analysis *Primary Mental Abilities* (1938). They are: (i) *Completion*; (ii) *Grammar*; and (iii) *Controlled Association*. Completion and Grammar were found to have high loadings (.83 and .80 respectively) and Controlled Association a moderately high loading (.66) on the general verbal fluency factor isolated by Rogers (1953) in his study of verbal fluency.

The spatial tests were: (i) the *National Institute of Industrial Psychology Group Test 81*; (ii) the *Second Peel Group Test of Practical Ability*; and (iii) *Blocking Counting* (Thurstone). Block Counting was also used by Thurstone in his P.M.A. battery with a specified time limit of 7 minutes. The results of preliminary testing, however, showed this time to be insufficient for the younger groups now tested, and accordingly 9 minutes was allowed for the test in the main experiment.

The four number tests from Thurstone's P.M.A. battery were used, namely *Addition*, *Subtraction*, *Multiplication*, and *Division*. Thurstone maintains that simple number tests are best for identifying a number factor, and the same view is held by Coombs (1941). Wrigley (1956), in his analysis of ability in elementary mathematics also used the Thurstone tests to identify a number factor. The tests are highly speeded, so much so that the results of preliminary testing suggested that the time limit (3 minutes for each test) be extended for the main experiment. Four minutes for each test was then allowed. Also, to avoid an undue influence being exerted by the number tests in the analysis—the number tests after all occupied far less testing time than the tests from any other group—the scores on Addition and Subtraction were combined. Only three number tests appear therefore in the battery for factorial analysis, namely (i) Addition and Subtraction, (ii) Multiplication, and (iii) Division.

#### IV.—DESCRIPTION OF GROUPS TESTED.

Testing was carried out in three co-educational grammar schools in Belfast. All pupils completing the Junior Certificate course in experimental science in these schools were tested. The numbers involved, together with the ages of each group at the time of testing, are shown in Table 2. In schools X and Z all the pupils were in their fourth school year. In school Y, on the other hand, some

TABLE 2  
AGES AND SIZES OF GROUPS.

School	Boys or Girls	Numbers	Average Age (yrs. : months)	Age Range (years : months)
X	B	47	15 : 4	14 : 8—16 : 2
	G	33	15 : 2	14 : 10—15 : 7
Y	B	76	14 : 10	13 : 5—16 : 1
	G	41	14 : 10	13 : 9—16 : 0
Z	B	69	15 : 7	14 : 11—16 : 6
	G	41	15 : 6	14 : 10—16 : 2

were completing the course in their third school year. All had been taught physics, chemistry, and biology as elements of a unified general science course. The smaller number of girls is due to the fact that all three schools allowed some of their academically weaker girls to take domestic science as an easier option.

#### V.—ANALYSIS OF RESULTS.

In each school the mean score on each of the tests was calculated for boys and girls separately. For both boys and girls the differences in mean score on all the science tests, and also on some of the other tests, were found to be significant at the 5 per cent. level. Correlations between tests were, therefore, derived for each school group (see Lindquist, 1940, 219-222), and averaged via Fisher's  $z$  transformation.\* The averaged correlations between all pairs of tests are shown in the Appendix. A centroid analysis then gave the factor loadings shown in the first section of Tables 3 A and B.

We see that for both boys and girls the first bipolar factor, Factor II, separates the three number tests from the rest. In the case of boys the second bipolar factor, Factor III, then separates the three science tests from the verbal and spatial tests, while the next factor, Factor IV, distinguishes between the verbal and spatial tests. In the case of the girls Factor III separates the verbal and science tests from the spatial tests, while Factor IV distinguishes between the verbal and science tests. The two analyses thus show a general agreement in indicating the presence of separate group factors (though possibly overlapping) for the verbal, number, spatial and scientific group of tests.

TABLE 3  
FACTOR LOADINGS.

#### A.—Boys.

Test	Centroid Factors					Rotated Factors					Communality
	I	II	III	IV	V	G	V	N	S	Sc	
Completion ...	635	154	206	-381	-092	486	593	120	079	114	0.621
Control'd Assoc.	588	088	152	-409	-063	400	574	162	-099	147	0.547
Grammar .....	628	027	247	-368	028	394	611	234	021	089	0.591
Add. and Sub.	330	-774	-068	044	-088	034	-097	841	-069	015	0.723
Mult. ....	307	-758	-060	-032	207	-134	041	818	126	102	0.715
Division .....	362	-680	128	-076	-110	092	100	768	-085	-110	0.628
N.I.I.P. Test ..	620	237	195	403	035	679	017	036	424	018	0.643
Peel Test ....	457	291	248	428	-084	643	-057	-081	328	-114	0.544
Blocks .....	422	348	319	403	156	529	055	-146	522	104	0.587
Physics .....	608	420	-464	068	175	472	059	-132	215	713	0.798
Chemistry ....	545	319	-569	057	130	359	063	-065	068	776	0.744
Biology .....	561	328	-335	085	-291	591	078	-069	-220	466	0.626
Sum of Squares	3.230	2.316	1.004	0.980	0.241	2.434	1.096	2.117	0.703	1.417	7.767
Contribution to Variance (per cent.) .....	26.9	19.3	8.4	8.2	2.0	20.3	9.1	17.6	5.9	11.8	64.7

\* The process is one in which the groups are regarded as samples drawn from the same population (or from equally correlated populations), the averaged values being estimates of the correlations for this population.



## B.—GIRLS.

Test	Centroid Factors					Rotated Factors					Communal- ity
	I	II	III	IV	V	G	V	N	S	Sc	
Completion . . .	483	265	318	-412	-041	449	603	-069	-066	029	0.575
Control'd Assoc.	480	184	306	-423	-057	416	602	006	-063	010	0.540
Grammar . . . .	599	161	234	-350	220	656	409	058	-084	-049	0.610
Add and Sub.	283	-696	087	040	179	132	-103	741	-162	-026	0.604
Mult. . . . .	369	-660	-029	-116	176	209	-011	740	-038	-154	0.617
Division . . . .	297	-615	-056	081	-028	035	-060	682	063	051	0.477
N.I.I.P. Test . .	634	107	-493	033	132	566	-080	130	574	-018	0.673
Peel Test . . . .	532	225	-513	060	148	528	-128	-017	571	-027	0.622
Blocks . . . . .	511	244	-530	-076	-101	377	087	-031	682	-032	0.622
Physics . . . . .	649	236	043	392	148	450	058	029	287	602	0.651
Chemistry . . . .	573	272	261	510	061	528	-067	-042	003	669	0.733
Biology . . . . .	570	277	373	248	-154	419	268	-039	-008	613	0.625
Sum of Squares	3.147	1.756	1.257	0.978	0.216	2.265	1.018	1.592	1.250	1.218	7.344
Contribution to Variance (per cent.) . . . . .	26.2	14.6	10.5	8.2	1.8	18.9	8.5	13.3	10.4	10.2	61.2

- N.B. (i) A decimal point should precede each entry in the body of the tables.  
(ii) Each loading of centroid factors II-V which exceeds twice its standard error is in heavy type.

The statistical significance of the successive centroid factors was investigated by the formula of Burt and Banks (1947) for the standard error of a factor loading. Adopting Vernon's suggestion (1950, 130) that a factor be accepted as statistically significant if about half its loadings exceed twice their standard errors we see that for each analysis Factors II, III and IV may be accepted as significant. Factor V on the other hand, cannot be accepted as significant. Its loadings are included merely to enable rotations to a general and four group factors to take place.

Rotations of the centroid factors were carried out by the method described by Burt (1950). The rotated factor loadings are shown in the second section of Tables 3 A and B. Although all the negative loadings are not eliminated, most of those remaining are numerically small, and may be regarded as sampling fluctuations from zero. Factor G is a general factor except in so far as its loadings on the number tests are small, and indeed negative in one instance. Partly for this reason G's contribution to the total variance is less than that of centroid factor I. The verbal factor V, number factor N and spatial factor S are clearly defined by the verbal, number and spatial tests, respectively. Also the scientific factor emerges as a distinct group factor with negligible overlap on the non-science tests. The contribution of this factor to the communalities of all three science tests in both analyses is substantial. Evidently attainment in elementary science depends both on the general factor and on the science attainment factor.

Of special importance are the loadings on the science tests of the verbal, number and spatial factors. Of these, the loadings on the physics test of the spatial factor appear fairly prominent in both analyses. Attainment in elementary physics—though not in chemistry and biology—seems to be influenced to some extent by spatial ability, quite apart from the general and



scientific factors. The only other loading on the science tests suggesting an overlap is that of the verbal factor on the biology test for the girls' analysis. In contrast the corresponding loading for the boys' analysis is near zero—this being the only marked difference in factorial pattern between boys and girls. The verbal factor loadings on the chemistry and physics tests in both analyses are negligibly small.

A somewhat surprising feature of these results is the consistently small loadings of the number factor on the science tests. Attainment in all three branches of school science shows no appreciable overlap with the number factor. In view, too, of the very limited extent to which the number tests contribute to the general factor G, it would seem that attainment in elementary science is independent of numerical ability—at any rate as assessed by the number tests of this investigation.

## VI.—DISCUSSION.

The existence of a separate group factor for scientific attainment has important implications for the teaching of science in schools. As has been stated, all the pupils tested had been taught physics, chemistry and biology as parts of a unified course of general science. In each of the schools concerned, junior-form teaching was organized on the assumption that it is desirable for science in its early stages to be taught and conceived as a whole. Many science teachers would not accept this. They would argue that physics, chemistry and biology are fundamentally different disciplines, and that an attempt to minimize their differences by teaching only general science is harmful (see, for example, Bradley, 1958). It is, of course, possible that if the investigation had been conducted in schools where physics, chemistry and biology had been taught from the beginning as separate subjects, and by different teachers, a scientific group factor would not have emerged so prominently. What has been demonstrated conclusively, however, is that when the teaching of junior-form science is organized on a unified basis, a prominent scientific group factor exists, and can be isolated by factorial techniques. If such a factor had not been established, then the case for a unified approach to science teaching would be seriously weakened.

It remains true, of course, that attainment in science is closely related to general intelligence. Thus, in Tables 3 A and B, the loadings on the science tests of the general factor are, for the most part, moderately high. And although this factor cannot be identified with the general cognitive factor  $g$ , the results do suggest that marked success in any branch of school science would be unlikely without a moderate or high fund of general intelligence.

Some pupils, of course, would still show an appreciable disparity of attainment between the different branches. But this would not be inconsistent with the factor patterns obtained. Apart from the effect of each test's specificity—necessarily ignored in the analysis—there are the differences resulting from the overlap with spatial and (for girls) verbal ability. Thus, a relatively superior, or inferior, attainment in physics could be (at least partially) an effect of the spatial factor. In the same way a noticeably different attainment in biology for girls might be explained by the verbal factor.

A further point is that if scientific attainment had been measured by school examinations, the variance from verbal and possibly other group factors would almost certainly have been greater. Thus, with essay-type answers variations in the accuracy and fluency of verbal expression would inevitably appear, and so augment the verbal factor. Again, if the test of examination required the use and interpretation of diagrams a spatial factor might in this way be brought in.



The decision to minimize, if not eliminate, these extraneous variations by the use of objective tests was made deliberately. It was intended from the outset that, if possible, only factors inherent in the learning of elementary science should enter into the analysis.\*

Again, the lack of dependence of attainment in physics and chemistry upon the verbal factor does not mean attainment in these subjects is unrelated to verbal ability as generally understood by teachers and others concerned with school education. The verbal factor isolated in a factor analysis using orthogonal reference frames is, by definition, independent of the factor of general ability or attainment, whereas verbal ability is generally thought of as depending on general ability—of being a compound, in fact, of the general and verbal factors. In this sense of the term pupils of high scientific attainment would tend to have above-average verbal ability because of the general factor.

Despite a similar distinction between the number factor and numerical ability, it is surprising that performance on the science tests, and in particular on the physics test, has no overlap with the number factor. An understanding of science—or at any rate physical science—is generally held to be closely akin to mathematical or numerical aptitude. The results of this study, however, show that a number factor defined by speeded tests of computation has no relation to attainment in elementary science.

Not the least important finding is the general similarity in factorial pattern between boys and girls. This should be regarded not merely as a negative result, but rather as one justifying the conclusion that the factorial nature of attainment in science does not differ greatly for boys and girls.

Finally, it should be noted that the science tests used in this research are tests of a 'low-level,' factual kind. Almost all the test items in fact would fall into the first two of Bloom's (1956) classification of educational goals, namely *knowledge* (the recall of specifics and universals in the field) and *comprehension* (the understanding of what is being communicated without necessarily applying it to other material). Analyses of other aspects of scientific ability, such as application, analysis (of elements, relationships, etc.) and evaluation (of information, conclusions, etc.), are now being undertaken, and it is hoped to report on these in due course.

\* It should be mentioned, however, that an extraneous factor introduced by using the particular objective tests of the present study is one resulting from the *form* of the tests. As all three tests consist of multiple-choice items followed by completion items, an artificial formal factor is introduced (see Vernon, 1950, 76-77). This has the effect of spuriously raising the variance from the scientific group factor. Probably, however, this effect is relatively small, and in any case the factorial pattern would not be affected.

## VII.—REFERENCES.

- ALEXANDER, W. P. (1935). *Intelligence concrete and abstract*. *Brit. J. Psychol. Monogr. Suppl.*, XIX.
- BERRIDGE, A. R. (1947). *A Factorial Analysis of Ability in School Physics*. Unpublished M.A. Thesis, University of London.
- BLOOM, B. S. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals*. Longmans, Green and Co.
- BRADLEY, J. (1958). The concept of school general science. *Studies in Education*, III, 1, 60-63.
- BURT, C. (1950). Group factor analysis. *Brit. J. Psychol. (Stat. Section)*, III, 40-75.
- BURT, C., and BANKS, C. (1947). A factor analysis of body measurements for British adult males. *Ann. Eugen.*, III, 238-256.
- CHARY, K. V. (1948). *A Statistical Analysis of the Performance of Boys and Girls in Science*. Unpublished M.A. Thesis, University of London.

- COOMBS, C. H. (1941). A factorial study of number ability. *Psychometrika*, VI, 161-189.
- JAMES, E. (1951). *An Investigation into the Factors entering into the Ability to learn Biology at School Certificate Stage*. Unpublished M.A. Thesis, University of London.
- JOG, R. N. (1955). *A Study of the Factors underlying the Ability of Grammar School Boys at the School Certificate Stage to learn Physics*. Unpublished M.A. Thesis, University of London.
- LAMB, F. C. (1954). *A Factorial Analysis of Ability in School Biology*. Unpublished M.Ed. Thesis, University of Nottingham.
- LEWIS, D. G. (1961). Group factors of attainment in grammar school subjects. *Brit. J. Educ. Psychol.*, XXXI, 241-248.
- LINDQUIST, E. F. (1940). *Statistical Analysis in Educational Research*. Houghton Mifflin Co.
- ORMISTON, M. (1939). The bearing of general and special abilities upon scholastic success at the beginning and end of a secondary school career. *Brit. J. Educ. Psychol.*, IX, 164-173 and 213-223.
- PAWLEY, G. C. (1937). *A Factor Analysis of the Abilities involved in learning School Chemistry*. Unpublished M.A. Thesis, University of London.
- ROGERS, C. A. (1953). The structure of verbal fluency. *Brit. J. Psychol.*, XLIV, 4, 368-380.
- THURSTONE, L. L. (1938). *Primary Mental Abilities*. Chicago : University of Chicago Press.
- VERNON, P. E. (1950). *The Structure of Human Abilities*. London : Methuen.
- WRIGLEY, J. (1956). *The Factorial Nature of Ability in Elementary Mathematics*. Unpublished Ph.D. Thesis, Queen's University of Belfast.

(Manuscript received 12th March, 1963.)



## VIII.—APPENDIX.

TABLE 4

CORRELATIONS BETWEEN TESTS.

## A.—BOYS.

1	2	3	4	5	6	7	8	9	10	11	12
Completion	552	584	022	077	195	309	199	291	342	285	388
2											
Contr. Assoc.	572	096	120	177	272	183	149	262	275	362	
3											
Grammar	152	159	256	303	178	216	257	244	283		
4											
Add. and Sub.	689	638	041	-084	-102	-127	-058	019			
5											
Mult.	609	000	-099	-169	-040	008	-177				
6											
Division	053	-005	-077	-156	-110	-006					
7											
N.I.I.P. Test	576	573	415	299	300						
8											
Peel Test	493	336	173	243							
9											
Blocks	302	155	166								
10											
Physics	718	619									
11											
Chemistry	586										
12											
Biology											

## B.—GIRLS.

1	2	3	4	5	6	7	8	9	10	11	12
Completion	540	567	-018	011	-057	164	099	215	221	193	397
2											
Contr. Assoc.	544	-049	145	-003	152	089	116	213	225	357	
3											
Grammar	123	126	002	302	282	199	281	268	313		
4											
Add. and Sub.	581	506	085	-036	-122	003	065	002			
5											
Mult.	500	135	059	107	062	-110	019				
6											
Division	170	038	019	087	052	-025					
7											
N.I.I.P. Test	630	629	386	318	175						
8											
Peel Test	565	380	251	197							
9											
Blocks	373	217	172								
10											
Physics	659	545									
11											
Chemistry	629										
12											
Biology											

N.B.—A decimal point should precede each entry in the tables.

# PROMPTING AND FEEDBACK IN THE LEARNING, RETENTION, AND TRANSFER OF CONCEPTS

BY M. C. WITTROCK AND P. A. TWELKER

*(University of California, Los Angeles)*

**SUMMARY.** 184 college students were taught to decipher coded sentences. Two variables were studied: (a) Rule: given or not given before the learner responded; (b) Knowledge of correct response (KCR); given or not given. A  $2 \times 2$  linear hypothesis design was used to analyze the four main experimental groups. As hypothesized, KCR increased learning, retention, and transfer only when the rule was given. It was concluded that KCR can be beneficial when it is not redundant.

Subjects who received rules or KCR rated the learning test as more enjoyable than subjects who received both rules and KCR or neither rules nor KCR.

## I.—INTRODUCTION AND OUTLINE OF PREVIOUS RESEARCH.

A BASIC problem in educational psychology is the determination of when to introduce information into complex, verbal learning to achieve optimal retention and transfer. Much of the recent work in programmed instruction assumes that learners should be led to make many overt and correct responses and that then these learners should be given immediate reinforcement. The response variables are greatly emphasized. It is implied in this position that when there is not an initially high probability of a correct response, information must be given to a learner before he makes a response *and* that information or affectively loaded feedback must also be supplied after a learner has correctly responded. Almost no distinction appears to be made between information of a specific nature (e.g., an answer to a given problem) or of a more general nature (e.g., a rule). Neither is there much distinction made between different types of learning outcomes (e.g., denotative meaning, and connotative meaning, transfer, savings, and motivation).

In much of the recent literature on programmed instruction a distinction is maintained regarding when information is introduced. When information is supplied to learners before they make overt responses, it is often termed 'prompts.' When the information is supplied after the responses have been made, it is often called knowledge of correct response, KCR, or in more general terms, reinforcement. With one type of KCR, confirmation, the learners are told only that their answers are correct or are incorrect. Other types of KCR may include specification of the actual correct responses and sometimes specification of the rules which pertain to the responses. Confirmation has usually been found to be less effective for learning than the more explicit varieties of KCR (Hirsch, 1952; Irion and Briggs, 1957). These more detailed varieties of KCR and their relationships to comparable prompting procedures have not been extensively studied in complex verbal learning. In paired-associate learning, prompting has been found to be more effective than confirmation (Cook and Spitzer, 1960).

Close examination of some recent studies, on direction in human learning indicates that the presentation of KCR has not always been carefully controlled.



For example, Craig (1956) and Kittell (1957) provided confirmation to each subject by use of punchboards. The terms 'directed' and 'maximum direction' which were used to describe experimental treatments, pertained only to the information given before students' responses. Since all groups in both experiments were given confirmation, there was definite direction and information supplied even to the 'independent' or the 'minimum direction' groups.

In other studies which are often compared to the above two (Haslerud and Meyers, 1958; Kersh, 1958) no KCR was given to any of the groups. These latter two studies obviously differ from the former two in the amount of direction supplied after the learner has responded. Since there is evidence that KCR in its various forms is an important factor in learning (Michael and Maccoby, 1961; Hirsch, 1952; Irion and Briggs, 1957), the obtained effects attributed to prompting may be due in part to KCR.

In an earlier experiment (Wittrock, in press) the question of which type of information to introduce was studied with different types of prompts. As predicted, dependable differences in learning, retention, and transfer were found which indicated that although initial learning was greatest with the most directive procedure, retention and transfer were greatest when rules or answers, but not both, were given to the learners. Giving neither rules nor examples produced the least effective learning. KCR was not given to any of the groups. The results were explained in terms of mediation theory, and in terms of practice and reinforcement which were available as a result of the introduction or withholding of information.

The work begun in that study was continued in this experiment. The time at which information was introduced and the type of information introduced was related to cognitive and motivational outcomes.

It was hypothesized that KCR enhances learning, retention, and transfer when the information it contains is not greatly redundant (Hypothesis I). The hypothesis differs from the widely accepted point of view that learners should necessarily make few mistakes *and* should receive immediate reinforcement. The hypothesis assumes that in the learning of concepts it may be the initial presentation of the information rather than the redundant presentation which produces the larger increment in learning. The hypothesis is consistent with results from paired associate learning (Angell and Lumsdaine, 1961).

The relationship between subjects' affectivity ratings and the information provided to the subjects was also studied. Feedback can have effects upon affectivity as well as upon cognition. It was hypothesized that subjects who received KCR or prompts would rate the experiment as more enjoyable than subjects who did not receive prompts (Hypothesis II). The second hypothesis was contingent upon the first and could be tested only if the first hypothesis was verified. The second hypothesis assumes that prompts or KCR help to increase the number of correctly deciphered sentences and thus provide added reinforcement.

## II.—PLAN OF THE INVESTIGATION.

(i) *Subjects.* The subjects were 184 University of California, Los Angeles, undergraduates from two upper division educational psychology courses taught by the senior author. Participation in the experiment was part of the laboratory requirement for the courses. There were twenty-nine males and 155 females. Means and standard deviations for age, verbal ability, and number ability are

shown by experimental treatment group in Table 1.

TABLE 1

MEANS AND STANDARD DEVIATIONS FOR THE AGE IN YEARS, VERBAL SCORE, AND NUMBER ABILITY MEASURES OF THE TREATMENT GROUPS.

Groups	Age	Co-op. English Score	Arithmetic Concepts Score
1. Rule not given, KCR not given . . . .	21.39 (2.46)	63.30 (6.00)	35.13 (6.09)
2. Rule not given, KCR given . . . . .	22.10 (5.06)	61.81 (9.92)	34.03 (8.87)
3. Rule given, KCR not given . . . . .	21.18 (1.72)	64.61 (7.83)	31.30 (7.40)
4. Rule given, KCR given . . . . .	21.73 (2.61)	63.19 (7.82)	33.24 (7.02)
5. Control A : Rule given and Answer given . . . . .	22.42 (4.94)	63.00 (6.09)	34.29 (8.02)
6. Control B. : No Learning Experience..	22.90 (6.74)	63.76 (7.19)	33.97 (7.80)

NOTE.—The Co-operative English Test, form S (1950) was used to measure verbal ability. A measure of number ability was obtained from the Arithmetic Concepts Test, a University of California, Los Angeles product which has not been published. (University of California, 1950.)

(ii) *Experimental Design.* Four groups were formed on the basis of two variables : (a) Rule—given or not given ; (b) KCR—given or not given. The four experimental groups were :

- (1) Rule not given, KCR not given ( $R_{ng} K_{ng}$ ).
- (2) Rule not given, KCR given ( $R_{ng} K_g$ ).
- (3) Rule given, KCR not given ( $R_g K_{ng}$ ).
- (4) Rule given, KCR given ( $R_g K_g$ ).

In addition, two control groups were included :

- (5) Control A : Rule given and Answer given, KCR not given ( $R_g A_g K_{ng}$ ).
- (6) Control B : No Learning Experience.

Subjects were individually assigned at random to each of the six treatments.

(iii) *Materials.* The problems for each group were presented in programmed booklets. The booklets were of 24 weight, gray, opaque paper, and were  $8\frac{1}{2}$  inches wide, and  $3\frac{3}{4}$  inches high. They were fifteen pages long for groups without KCR and twenty-six pages long for groups with KCR. The differences in thickness were not readily apparent. The first two pages contained general instructions. The third page contained a detailed explanation of the task with an appropriate example. For the KCR groups, the next page gave an example of an enciphered sentence, the deciphered sentence, and the appropriate rule. The last page of directions summarized the previous pages. The problem solving task consisted of ten enciphered sentences. Each group was to decipher examples



of transpositional cryptograms. In addition, a rule was given to groups 3, 4 and 5; the answer to the cryptogram given to group 5; and KCR given to groups 2 and 4. To hold overt practice constant across treatments, subjects were required either to derive a rule or to copy the prompted rules in spaces provided in the booklets. Every booklet contained the ten items mentioned above, but the items were presented in two different sequences to control for position effects. Each sentence contained eighteen letters transposed according to a systematic rule. When properly deciphered the sentences were complete and sensible. An example of a rule, enciphered sentence, and deciphered sentence appear below:

Rule: Put all the letters in the coded sentence in reverse order; then separate these letters into words.

Enciphered sentence: TAHE KILK LATNA COHW.

Deciphered sentence: WHO CAN TALK LIKE THAT.

The learning test consisted of ten items with appropriate instructions. Each of the ten rules previously learned was represented by one item. The items were arranged in random order.

The retention and transfer test consisted of twenty-four enciphered sentences. Eight sentences represented the previously learned examples, eight represented new examples of previously learned rules, and eight represented new examples of new rules. The three types of sentences were presented in random order. Two different forms of the test were used to control for position effects. A sentence was marked as incorrect if more than two letters were misplaced.

(iv) *Procedure.* One week before Day 1, subjects were told of the experiment and times were arranged for them to report to the laboratory. When they reported, they were given the materials appropriate to their experimental treatment. The subjects in the KCR groups were asked especially to note the instructions for the KCR page. These subjects were told that the page containing the KCR was to be studied only after they had attempted to decipher the sentence, and that no credit would be given if this instruction was not followed.

The different treatments existed simultaneously in the same room. Subjects worked individually and were proctored by the experimenters. The learning test was given immediately after the subjects finished the ten problems. The retention and transfer test was given three weeks after the learning test. In all cases subjects were allowed as much time as they required to attempt to solve each item. After completion of each of the above two tests, each subject rated the test on a seven-point affectivity scale.

Each subject was told at the time of the post-test that he would be informed of the results of the experiment. A personal letter was mailed to each subject as promised.

### III.—STATEMENT OF RESULTS\*

The general linear hypothesis model (Kempthorne, 1952, pp. 38-67) and analysis of variance were used to analyze the data. After appropriate F tests and graphs of frequency distributions were examined, homogeneity of variance and normality of distribution were assumed for all criterion measures except one. These assumptions were violated for the problems score. However, with

\* The authors wish to thank the Western Data Processing Center for the use of the IBM 7090 computer in the analysis of the data.

an  $F$  of 244.90 and with the sizeable mean difference between the Rule given and Rule not given groups, there is little doubt that these results are dependable. Table 2 presents a summary of the results of the problems score on the learning test, and of the three-week retention and transfer tests. Table 3 presents a summary of the time required to learn and to perform, and of the affectivity ratings.

TABLE 2

TREATMENT GROUP MEANS AND STANDARD DEVIATIONS FOR THE LEARNING TASK SCORES, LEARNING TEST SCORES, AND THREE-WEEK RETENTION AND TRANSFER TEST SCORES.

Groups	Learning		Three-Week Retention and Transfer				
	Prob- lems	Test	Reten- tion Test (A)	Transfer to New Exam- ples Test (B)	Transfer to New Rules and New Exam- ples (C)	Com- bined Transfer Test (B and C)	Com- bined Reten- tion and Transfer (A and B and C)
1. Rule not given, KCR not given, $n=23$ ....	2.52 (2.27)	2.65 (2.35)	3.13 (1.84)	2.87 (2.03)	2.48 (1.56)	5.35 (3.35)	8.48 (5.02)
2. Rule not given, KCR given, $n=31$ .....	5.16 (2.92)	5.45 (2.53)	5.48 (2.11)	4.90 (1.87)	3.52 (2.01)	8.42 (3.55)	13.90 (5.20)
3. Rule given, KCR not given, $n=33$ .....	9.52 (1.00)	8.06 (1.90)	5.79 (1.69)	5.30 (1.53)	3.76 (1.77)	9.06 (2.82)	14.85 (4.14)
4. Rule given, KCR given, $n=37$ .....	9.43 (1.34)	7.43 (2.41)	5.70 (1.85)	5.35 (1.78)	3.35 (1.38)	8.70 (2.70)	14.41 (4.20)
5. Control A: Rule given and Answer given $n=31$ .....	10.00 (0)	7.90 (2.48)	5.55 (2.00)	5.10 (2.04)	3.26 (1.84)	8.35 (3.47)	13.90 (5.19)
6. Control B: No Learn- ing Experience, $n=29$		3.59 (2.64)	2.69 (2.17)	3.17 (1.95)	2.72 (2.07)	5.90 (3.76)	8.66 (5.83)

NOTE.—There were ten problems in the learning task and another ten in the learning test; on the three-week retention and transfer tests there was a total of twenty-four items, with eight items on each of the three sub-tests. Possible score on each of these tests was the respective number of items on the test. All columns list the mean number of correct items.

*The Problems Test.* From Table 4, the Rule factor produced a reliable effect upon learning ( $F=244.90$ ,  $p < .001$ ). The Rule given groups correctly deciphered more sentences than did the Rule not given groups. The KCR factor produced a reliable effect upon learning ( $F=12.62$ ,  $p < .001$ ). There was also a reliable interaction ( $F=14.30$ ,  $p < .001$ ) between the Rule factor and the KCR factor. The  $R_{ng} K_g$  group performed better than the  $R_{ng} K_{ng}$  group ( $t=3.73$ ,  $p < .001$ ). The difference between the two Rule given groups was negligible.

When time to learn was compared, the Rule factor produced a reliable effect ( $F=146.21$ ,  $p < .001$ ), as shown by Table 5. The Rule given groups took less time than the Rule not given groups.

Control group A outperformed the  $R_g K_g$  group ( $F=5.51$ ,  $p < .05$ ). When time to learn was compared, Control group A took less time than the  $R_g K_g$  group ( $F=25.25$ ,  $p < .001$ ).



TABLE 3

TREATMENT GROUP MEANS AND STANDARD DEVIATIONS FOR THE TIME IN MINUTES TO LEARN AND TO PERFORM, AND FOR THE AFFECTIVITY RATINGS.

Group	Time		Affectivity		
	Problems	Learning Test	Retention and Transfer Test	Learning Test	Retention and Transfer Test
1. Rule not given, KCR not given, n=23 . . . .	85.91 (26.27)	40.17 (21.45)	84.17 (30.76)	2.04 (1.07)	2.74 (1.14)
2. Rule not given, KCR given, n=31 . . . . .	84.81 (27.75)	49.74 (30.55)	91.97 (32.86)	3.06 (1.48)	3.55 (1.50)
3. Rule given, KCR not given, n=33 . . . . .	35.33 (8.75)	52.97 (22.10)	86.03 (26.33)	4.76 (1.25)	4.27 (1.40)
4. Rule given, KCR given, n=37 . . . . .	46.14 (15.27)	51.19 (16.78)	92.49 (31.57)	3.81 (1.54)	3.92 (1.44)
5. Control A: Rule given and Answer given, n=31 . . . . .	31.16 (7.04)	57.19 (19.72)	89.65 (27.95)	4.42 (1.77)	3.55 (1.48)
6. Control B: No Learning Experience, n=29 . . . .	—	77.83 (32.80)	75.48 (26.22)	2.86 (1.22)	3.24 (1.27)

NOTE.—The attitude measure was a seven-point scale from 1=Very frustrating to 7=Very enjoyable.

TABLE 4

SUMMARY OF THE F RATIOS FOR THE LEARNING TASK AND LEARNING, RETENTION, AND TRANSFER TESTS SCORES OF THE EXPERIMENTAL GROUPS.

Stimulus Variables	Learning		Three-Week Retention and Transfer Test				
	Task	Test	Retention (A)	Transfer to New Examples Test (B)	Transfer to New Rules and New Examples (C)	Combined Transfer Test (B+C)	Combined Retention and Transfer (A+B+C)
Rule . . . . .	244.90†	77.20†	17.61†	19.46†	3.26	12.63†	16.71†
KCR . . . . .	12.62†	6.67*	10.95†	10.16	1.05	5.82*	8.78†
Interaction . . . . .	14.30†	16.61†	12.66†	9.24†	5.47*	9.30†	12.18†

\* Significant at .05 level, df 1,120.

† Significant at .01 level, df 1,120.

‡ Significant at .001 level, df 1,120.

TABLE 5

SUMMARY OF THE F RATIOS FOR THE TIMES AND AFFECTIVITY RATINGS OF THE EXPERIMENTAL GROUPS.

Stimulus Variables	Time		Attitude		
	Learning Task	Learning Test	Retention and Transfer Test	Learning Test	Retention and Transfer Test
Rule .....	146.21 <sup>†</sup>	2.86	.05	46.66 <sup>†</sup>	14.01 <sup>†</sup>
KCR .....	1.73	.86	1.65	.02	.80
Interaction .....	2.60	1.82	.01	15.41 <sup>‡</sup>	5.23*

\* Significant at .05 level, df 1,120.

<sup>†</sup>—Significant at .01 level, df 1,120.

<sup>‡</sup> Significant at .001 level, df 1,120.

*The Learning Test.* The results of the test, given immediately after the problems had been worked, indicated that the Rule variable produced reliable differences in the scores ( $F=77.20$ ,  $p < .001$ ). The groups that were given the rule correctly deciphered more sentences than did the groups that were not given the rule.

The KCR factor produced a dependable effect upon the test scores ( $F=6.67$ ,  $p < .05$ ). There was also a reliable interaction ( $F=16.61$ ,  $p < .001$ ) between the two variables. The  $R_{ng} K_g$  group performed better than the  $R_{ng} K_{ng}$  group ( $t=4.18$ ,  $p < .001$ ). There was no dependable difference between the Rule given groups. Thus, the first hypothesis was verified.

No dependable differences were found between any two of the experimental groups when time to learn was compared. When the affectivity ratings were compared, the Rule factor produced reliable differences ( $F=47.66$ ,  $p < .001$ ). The Rule given groups rated the test as more enjoyable than the Rule not given groups, as shown by Table 3. The KCR variable produced no dependable effect on activity rating, although an interaction between the Rule and KCR factors was statistically reliable ( $F=15.41$ ,  $p < .001$ ).

The  $R_{ng} K_g$  group rated the task as more enjoyable than the  $R_{ng} K_{ng}$  group ( $t=2.91$ ,  $p < .01$ ). However, the  $R_g K_g$  group rated the task as less enjoyable than the  $R_g K_{ng}$  group ( $t=2.88$ ,  $p < .01$ ). The presentation of KCR increased affectivity when learners were not given rules but decreased affectivity when learners were given rules.

There were no reliable differences between Control group A and the  $R_g K_g$  group on the test score, the time required for the test, or the affectivity rating.

The odd-even Spearman-Brown coefficient of reliability for this test was reported by Wittrock (in press) to be .92.

*The Retention and Transfer Test.* The Rule and KCR factors produced reliable effects on the retention subtest and on the transfer to new examples subtest. The interaction was reliable for the retention subtest and both transfer subtests. The  $R_{ng} K_g$  group outperformed the  $R_{ng} K_{ng}$  group on the retention subtest ( $t=4.35$ ,  $p < .001$ ), on the transfer to new examples subtest



( $t=3.76$ ,  $p<.001$ ), and on the transfer to new rules and new examples subtest ( $t=2.12$ ,  $p<.05$ ). There were no dependable differences between the Rule given groups or between the  $R_{ng}$   $K_{ng}$  group and the Rule prompted groups. The results cited above for the subtests also apply to the combined transfer test and to the combined retention and transfer tests (see Tables 2 and 4). Thus, the first hypothesis was verified.

When time to perform on the combined retention and transfer test was considered, there were no reliable differences among the groups. There were, however, reliable differences for the affectivity ratings. The Rule factor produced a reliable effect upon the affectivity rating ( $F=14.01$ ,  $p<.001$ ). Both Rule given groups rated the task as more enjoyable than the Rule not given groups. Hypothesis 2 was supported. The interaction between Rule and KCR factors was also dependable ( $F=5.23$ ,  $p<.05$ ). The  $R_{ng}$   $K_g$  group rated the task as more enjoyable than the  $R_{ng}$   $K_{ng}$  group ( $t=2.25$ ,  $p<.05$ ). There was no dependable difference between the Rule given groups.

The odd-even Spearman-Brown coefficient for combined treatments was .81 for the retention subtest, .76 for the transfer to new examples subtest, and .60 for the transfer to new principles subtest. For the combined transfer subtests, the coefficient was .82 and for the combined retention and transfer subtests, .91.

#### IV.—DISCUSSION AND CONCLUSIONS.

The first hypothesis was supported. It stated that prompts or non-redundant KCR, would enhance learning. The prompted rules produced a reliable effect on the initial problems, learning, retention, and combined transfer tests, and on time to learn. The KCR factor and rule plus KCR interaction produced a reliable effect on these same criteria except on the time to learn criterion. From the above primary and interactional effects, it appears that non-redundant information in the form of KCR added to a minimally directed situation enhanced learning while redundant KCR added to an already prompted situation, did little to learning, retention, and transfer. However, the redundant information did not appear to reduce learning, retention, and transfer. This study lent support to those who found that prompting is more effective than lack of direction (e.g., Wittrock, in press; Kittell, 1957). Prompting compared with less direction can also reduce the amount of time required to learn.

Control group B, No Learning Experience, performed about as well on the post-test as did the minimally directed  $R_{ng}$   $K_{ng}$  group. One implication is that, under certain conditions, practice with very little direction can be an ineffective way to learn.

Data partially supported the second hypothesis. Prompts produced correct answers and positive affectivity ratings. Curiously, KCR produced no primary effect on the affectivity measure. The interactional effect between prompts and KCR indicated that KCR increased positive affectivity in a minimally directed situation, but decreased positive affectivity in a prompted situation. Prompts without KCR produced the most positive rating, while  $R_{ng}$   $K_{ng}$  produced the most negative rating. When minimally directed situations are frustrating to the subjects, this frustration may not be completely negated by the presentation of KCR. It would also seem that KCR plus prompts may tend to be overly redundant or directive.

The data of this study are interpreted to indicate that, in minimally directed situations comparable to the ones sampled in this study, learning can be enhanced with the introduction of prompts or KCR, but that introduction of further redundant information may lower a positive affectivity rating of a learner.

#### V.—REFERENCES.

- ANGELL, D., and LUMSDAINE, A. A. 1961. The effects of prompting trials and partial-correction procedures on learning by anticipation. San Mateo, Calif.: *American Institute for Research*, AIR—C14-9/61SR5; AFOSR—1343.
- COOK, J. O., and SPITZER, M. I. 1960. Supplementary report: Prompting versus confirmation in paired-associate learning. *J. Exp. Psychol.*, 59, 275-276.
- CO-OPERATIVE TEST DIVISION. 1950. *Co-operative English Tests, Form S*. Princeton, N.J.: Educational Testing Service.
- CRAIG, R. C. 1956. Directed versus independent discovery of established relations. *J. Educ. Psychol.*, 47, 223-234.
- FORGUS, R. H., and SCHWARTZ, 1957, R. J. Efficient retention and transfer as affected by learning method. *J. Psychol.*, 43, 135-139.
- HASLERUD, G. M., and MEYERS, SHIRLEY, 1958. The transfer value of given and individually derived principles. *J. Educ. Psychol.*, 49, 293-298.
- HIRSCH, R. S. 1950. The effects of knowledge of test results on learning of meaningful material. (Pennsylvania State Univ. Instructional Film Research Program), Port Washington, N.Y.: U.S. Naval Training Devices Center, Office of Naval Research, Technical Report. S.D.C., 269-7-19.
- IRION, A. L., and BRIGGS, L. J. 1957. Learning task and mode of operation variables in the use of the subject-matter trainer. Air Force Personnel and Training Research Center, Technical Report AFPTRC—TR-57-8. (ASTIA Doc. No. AD134252).
- KEMPTHORNE, O. 1952. *The Design and Analysis of Experiments*. N.Y.: John Wiley and Sons, Inc.
- KERSH, B. Y. 1958. The adequacy of 'meaning' as an explanation for the superiority of learning by independent discovery. *J. Educ. Psychol.*, 49, 282-292.
- KITTELL, J. E. 1957. An experimental study of the effect of external direction during learning on transfer and retention of principles. *J. Educ. Psychol.*, 48, 391-405.
- MICHAEL, D. N., and MACCOBY, N. 1953. Factors influencing verbal learning from films under conditions of audience participation. *J. Exp. Psychol.*, 46, 411-418.
- WITTROCK, M. C. 1962. Mediation theory applied to discovery learning: Cueing and prompting mediated responses. Unpublished manuscript, Education Library, University of California, Los Angeles.

(Manuscript received January 30th, 1963.)



# THE USE OF GROUP TESTS OF ABILITY AND ATTAINMENT : A FOLLOW-UP STUDY FROM PRIMARY TO SECONDARY SCHOOL

By NORMAN FRANCE

(*Chief Educational Psychologist, Kent Education Committee*)

**SUMMARY.** This paper describes a six-year follow-up study of the 1,300 children in a random (1 in 15) sample of all Kent children born in 1945. They entered their secondary schools in September, 1956, and were tested in December, 1959. Information concerning them included scores on certain group tests of ability and attainment which were given in 1953, 1954, 1955 and 1956, and assessments made by primary school teachers in 1956 and by secondary school teachers in 1959. Selection for secondary education in Kent since 1956, has been largely based on the cumulative school record card and has incorporated refinements designed to reduce competition between primary schools and to ensure equal opportunity for all children in Kent. The results of this analysis suggest that a selection procedure based on the result of a series of group tests taken throughout the junior school course and incorporating both primary school assessments and an adjustment for differences between primary schools can provide an accurate forecast of later secondary school achievement. The correlation between the primary and secondary school criteria for the unselected sample was 0.91 for an interval of nearly four years.

## I.—INTRODUCTION.

GROUP tests of ability and attainment have been in use in Kent in connection with the cumulative school record card for nearly forty years; their introduction is recorded in the *Kent Education Gazette* of October, 1923. The early tests were those of Burt and Ballard. These were later replaced with tests specially constructed for use in Kent schools. The first of these was the Spelling Test prepared by Vickers (Inspector of Education for the Kent Education Committee) and introduced into Kent schools in 1938. By the middle fifties there existed three parallel batteries of Kent tests (A, B and C) to be administered in junior schools in consecutive years. Each battery consisted of verbal and non-verbal intelligence tests, three tests of English and two of Arithmetic. These tests were intended to be given in place of a school examination in each Summer term. Some of the earlier tests are now being replaced by new tests devised by members of the Kent School Psychological Service.

From the early days it had been intended that the information provided by the tests should be used as the basis for educational guidance at all levels of ability. In particular, the use of the record card to decide "the course of post-primary education most suited to a child's needs" was recommended in the 1930 edition of "Notes on the Record Card." It was claimed that this would be "a safer guide than a one-day examination, however skilfully conducted."

## II.—OUTLINE OF THE FOLLOW-UP INVESTIGATION.

Pointers to the future of the Kent selection procedures were firmly established when the Kent Education Committee decided, in 1955, to give yet more weight to the cumulative school record. The evolutionary process has now moved a long way towards complete reliance on this. Local selection panels now make their recommendations almost entirely on the evidence of the

individual record card. Even further simplification is envisaged and these changes together with a shift of emphasis in the aims of the three different types of secondary school will go a long way towards removing what is often a source of tension for parents, teachers and children.

This follow-up investigation is concerned with some evaluation of this process of revision of the selection procedure. When changes take place it is necessary to examine impartially all evidence of both old and reformed ideas. The children used in this investigation entered their respective secondary schools in 1956. The follow-up in December, 1959, was intended to assess children of all levels of ability before they achieved the age of 15. The tests employed comprised the Moray House Adult Intelligence Test I, the Manchester Reading Comprehension Test and a Kent Mathematics Test, and all were re-standardised on the Kent population. Heads of all secondary schools were asked to complete questionnaires in respect of each child in the sample and in particular to provide their own estimates of present ability and attainment and of corresponding future potential.

For each child, eighty items of information were collected and recorded, covering many aspects of primary and secondary school work. Fifty items referred to details of tests in the junior school and of supplementary tests taken at the end of the junior school course and the remaining items summarised the information on the secondary head's questionnaire completed after  $3\frac{1}{2}$  years in the secondary school.

Side by side with the development of school records, there have been changes in the selection procedures for transfer from primary to secondary education. In common with many other counties, Kent children are given intelligence and attainment tests. However, certain features are peculiar to Kent.

Whilst the local selection panels have complete discretion to allocate children to secondary school courses on the basis of the information on the record card, they receive some guidance in the shape of a 'combined score' based on fourth year tests. This combined score enables the panel to make a rough comparison between individual children. This combined score is calculated at Head Office for each school separately as the equally weighted sum of three items:

- (a) The standardised intelligence score (verbal+non-verbal).
- (b) The standardised educational rating (the educational rating is the mean of the attainment test scores: a further calculation gives the standardised educational rating. This ensures that the school's average attainment is adjusted to the level of the school's average intelligence score).
- (c) The standardised order of recommendation (the Head's order of recommendation is quantified school by school on the basis of the intelligence test score).

The intelligence test is taken under carefully controlled conditions after adequate teaching and practice, so that it may be used as a reliable base for (a) the adjustment to offset the school's own general influence on attainment and (b) the quantification of the order of recommendation. These adjustments do much to offset the undesirable competition between primary schools which is often based on selection procedure results.

The selection at 11+ is fully supported by a testing programme at ages 8, 9, 10 and 11, and this is peculiar to Kent.



## III.—SPECIAL FEATURES OF THIS FOLLOW-UP INVESTIGATION.

(a) *The Sample.*

The sample was composed at all stages of the children in schools (private and maintained) who were born on the fifth or twentieth day of any month of 1945. Every effort was made to complete the assessment of all these children at each stage, but inevitably there proved to be some variation in the composition of the sample at various stages.

The following table shows the size of the sample in 1956 and in 1959 together with the numbers in the Kent child population :

TABLE 1  
COMPOSITION OF THE KENT TWO-DAY SAMPLE.

	Group	Size of Kent child population		Size of Two-day sample (Percentage of corresponding population)	
		Boys	Girls	Boys	Girls
Jan., 1956	In Primary schools :	11,457	11,027	667 (5.8%)	611 (5.5%)
Dec., 1959	In Secondary schools :				
	Grammar .....	1,745	1,514	123 (7.0%)	109 (7.2%)
	Technical High .....	973	1,145	72 (7.4%)	70 (6.1%)
	Modern .....	7,549	6,879	475 (6.3%)	446 (6.5%)
	All Secondary schools : ....	10,267	9,538	670 (6.5%)	625 (6.5%)

(b) *Standardisation of the Tests.*

In recent years a large number of tests have been standardised in Kent on single and several age-groups. A technique has been developed of effecting these standardisations on the principle of equating percentile levels to standard scores using a graphical method. At first Arithmetic Probability graph paper was used but recently it has been found more convenient to convert percentile levels to standard scores by means of a specially prepared table and to plot raw score against 'true' standard score level on ordinary graph paper. The feature of this kind of treatment is that a normal distribution is represented by a straight line whose gradient is proportional to the standard deviation. Thus, the 'geometry' of the straight line (determined by a point in the plane and the gradient) matches the 'geometry' of a normal distribution (determined by the median score and the standard deviation). These representations provide a powerful instrument for the visual comparison of two or more distributions.

(c) *Secondary School Criteria.*

The main problem in follow-up investigations is to decide on the criterion of secondary school 'success' and thus on the extent to which the whole range of ability can be explored in different types of secondary school. With

8-408

8970

these considerations in mind, it was decided to consider as many criteria as possible which could be regarded as relevant. The following items were assembled in December, 1959 (when the children were aged 14) :

- (1) Scores on test battery ( $M+E+I$ ).
- (2) Secondary school Head's assessments ( $M+E+I$ ).
- (3) Forecast of total number of O level passes.
- (4) Forecast of total number of A level passes.
- (5) Forecast of type of further education.
- (6) Forecast of length of secondary school life.
- (7) The answer to the question " Was the allocation in 1956 subsequently justified? Yes/No."

Examination of these criteria was reckoned as an essential part of the investigation.

A special feature of the secondary school Head's assessments was the request to Heads to give their estimates in the form of standard scores ( $M=100$ ,  $s.d.=15$ ) This is the form in which Kent test scores are entered on the primary school record card. For this reason, Heads of secondary schools are familiar with this way of indicating percentile level in the age-group.

(d) *Comparisons of Primary and Secondary School Assessments.*

The relationship between primary and secondary school success was also examined in several ways. Firstly, the amount of overlap was calculated and compared for the best in the modern schools and the weakest in the grammar and technical high schools for each single criterion and for a combination of two or more criteria. Secondly, the quality of children in these overlap zones was examined by means of a ' follow-back ' investigation for children in the two border-zones, as determined (a) by the Head's subjective judgment and (b) by the test battery. Thirdly, the border-zone groups at age 11 were followed up and compared on the basis of the age 14 criteria. Fourthly and finally, a correlational study related selected items at ages 8, 9, 10 and 11 with items at age 14.

(e) *Computational Aids.*

The Hollerith punched card system was used to tabulate the eighty items of information about each of the 1,000 or so children in the sample. Thus, it was possible to sort the cards on various combinations of columns in turn in order to produce the tables of distribution preparatory to the calculation of product-moment coefficients of correlation. The latter were calculated with the assistance of the Department of Education, University of Manchester, and of the computer in the Electrical Engineering Laboratory.

#### IV.—RESULTS OF THE KENT FOLLOW-UP INVESTIGATION.

From a very detailed and complex study, the following specific results are of interest :

(i) *Amount of overlap between selective and non-selective schools.*

A border-line of 110 standard score on ( $M+E+I$ ) tests at 14, was chosen since this approximated to the division between selective and non-selective secondary school provision. (The mechanics of the analysis determined that



the measure must be a multiple of ten.) The two border-zones were so defined as to include all those in the selective schools who scored less than 110 and all those in the non-selective schools who scored 110 or more. The combined size of the two border-zones based on a criterion of standard score of 110 on the M+E+I tests was about 12 per cent. of the age-group.

The amount of overlap based on an appropriate border-line of one of the single criteria :

- (a) M+E+I tests
  - (b) M+E+I Head's assessments
  - (c) O level forecasts
  - (d) length of secondary school life
- averaged about 10 per cent. of the age-group.

The amount of overlap based on the combination of two of the criteria :  
 M+E+I tests and (a) M+E+I assessments,  
                                   or (b) O level forecasts,  
                                   or (c) Length of secondary school life  
 averaged less than 5 per cent. of the age-group.

In particular, the combination of the M+E+I tests and the Head's broad judgment as to whether allocation at 11 had been correct, produced an overlap of only 2 per cent. of the age-group.

(ii) *Characteristics of Children in the Overlap Zones.*

(a) *Overlap Zones determined by Head's subjective judgment.*

There were fifty-five children in the group determined by the subjective opinions of Heads. Thirty-nine of these children had been in the original border-zone at 11+ taking supplementary tests. For six more the Heads were not prepared to say definitely that they had been wrongly placed. This left four children in the grammar/technical high schools and six children in the modern schools (less than 1 per cent. of the whole sample) who, in the opinion of the Heads, had been wrongly placed at the age of 11.

The group of modern school improvers showed an average and overall gain of 5 points of standard score relative to the group of grammar/technical high school deteriorators, yet the forecast number of G.C.E. 'O' level passes was still greater for the deteriorator group than for the improver group. Apart from this, Head's assessments exaggerated the differences between improvers and deteriorators.

(b) *Overlap Zones determined by a standard score of 110 on M+E+I tests.*

In this second analysis the group of misfits as picked out by test scores at 14 was larger: forty-two modern school improvers and 110 grammar and technical high school deteriorators. The proportion of the age group in selective schools was nearer 30 per cent. than 25 per cent. and there were sex differences in test scores. Thus, the two border-zones were not symmetrical about the border-line. Out of the 152 children in this group of misfits, only eight children in grammar and technical schools showed an average test standard score of less than 100 and one child in modern schools a score of 120 or more. Hence, the serious misfits comprised (again) less than 1 per cent. of the age-group sample.

*(c) General Consideration.*

Of all the many items on the primary school record card, the most useful individual items for picking out future misfits were undoubtedly the reasoning arithmetic, comprehension, educational rating and the verbal intelligence scores. The least useful items were the spelling and non-verbal intelligence scores. For the most useful items the chances of a child becoming a misfit clearly depended on how much his standard score was above or below the arbitrary border-line of 110. For instance, a child with a standard score in reasoning arithmetic of 120 or more allocated to a modern school was more likely than not to be regarded later as an improver, and, similarly, the occasional child with a score of 100 allocated to a grammar school was more likely than not to become a deteriorator. At the stage of primary Heads' recommendations, it was possible to draw similar conclusions about the standardised educational rating, intelligence score and order of recommendation over the narrower range of 115 to 105.

Size of school had no significant effect on the production of misfits but there was a slight suggestion that modern school improvers tended to come from the larger primary schools.

Although the effect of age was not really significant, more of the misfits had birthdays in the July to September quarter than in the October to December quarter.

Apart from the Senior Entrance Examination at 13+ which, in 1958, still determined entry to some boys' technical high schools, there were ten transfers in the whole two-day sample between secondary schools during the 11+ to 14+ period and six of them were in this misfit group.

The differences of opportunity to take G.C.E. which exist between children in different types of secondary school is illustrated by a closer look at these groups. The modern school improvers were expected to obtain an average of just less than two passes at 'O' level, whereas the grammar and technical high school deteriorators were still expected to achieve an average of nearly three passes.\*

*(iii) Overlap of ability and attainment between modern and grammar/technical high schools.**Follow-up of 11+ Border-zones.*

The selection procedure in 1956 at the beginning of this evolutionary process had three stages in each of which a progressively narrowing border-zone was considered :

40 per cent. of the age group were the subject of the primary Head's order of recommendation ;

25 per cent. took Supplementary Tests ;

and 5 per cent. were interviewed by the Selection Panel.

(By 1962, however, the second and third stages had virtually been eliminated.)

Thus, the most outstanding children in 1956 were offered a grammar school place at Stage I, whereas the interview stage was reserved for the most doubtful cases.

\* To keep the problem in perspective it should be recorded that analysis showed 38 per cent. of all boys and 35 per cent. of all girls to be candidates for G.C.E. at 'O' level. Included in the proportions are 11 per cent. and 7 per cent., respectively, for boys and girls in secondary modern schools.



The subsequent performance at 14 of children in each of these border-zone groups was looked at in connection with the criteria listed above. In every case the group of the most outstanding children at 11 was still the group showing the highest scores at 14, and there was generally a gradation of scores on each of the criteria down to the group of lowest status at 11. The least differences were, perhaps naturally, revealed at the final border-zone of the interview stage. The group of children offered grammar school places showed some superiority (as measured by the criteria at 14) over the group of children offered modern school places, but the differences for any of the criteria were not significant except possibly for the secondary Head's assessment and his forecast of 'A' level success.

(iv) *Assessments at ages 8, 9, 10, 11, and 14.*

(a) *Sex differences.*

One important facet of this research was the analysis of the differences in attainment and ability between boys and girls at different stages of their school life. A universal feature in every test and in every assessment at any of the ages 8, 9, 10, 11 or 14 was the greater spread of scores for boys than for girls. This means, of course, that if the score of the average boy and of the average girl is the same then there are more high-scorers and more low-scorers amongst the boys than amongst the girls. Differences in the scores at the middle level of ability for boys and girls vary appreciably from test to test and mask this effect of differing spreads. Hence it is impossible to generalise about sex differences in attainment and ability over the whole ability range. However, if the age group is thought of as divided into three equal ability groups, the following results can be picked out from this analysis in the junior school. In the upper group there was generally little difference between boys and girls except for the slight superiority of boys for reasoning arithmetic, the educational rating, verbal and non-verbal intelligence. In the middle group girls were generally superior to boys except for verbal and non-verbal intelligence. In the lower group, girls were superior to boys in every test.

At the secondary school stage, the same feature of a greater spread of 14+ test scores and assessments for boys was very much in evidence. The differences in mathematics scores between the average boy and girl were, however, now so great that boys were still better than girls even in the lowest third of the age-group. The comprehension test results showed a similar but modified distribution in that there was little difference between boys and girls at the very lowest level, whereas the verbal intelligence test showed no sex differences in the middle ranges of ability.

Although Head's assessments were made independently of test scores, they still showed a larger spread for boys. The sex-difference for Heads' assessments at any level of ability was, however, slight, and very much less than that shown by the test scores. It was interesting to notice in this connection that their assessments were slightly but clearly biased in favour of the older child in the age-group and that there was a tendency for the Head of a secondary modern school to under-estimate the attainment of his abler boys.

TABLE 2

COEFFICIENTS OF CORRELATION BETWEEN JUNIOR SCHOOL TESTS AND CORRESPONDING SECONDARY SCHOOL TESTS AND SECONDARY HEADS' ASSESSMENTS.  
(Number of children in the sample is given in brackets.)

See Appendix for meaning of abbreviations.

Date of Testing	Mathematics		English		Intelligence	
	AUTUMN, 1959 (Age 14)					
	Test	Assessment	Test	Assessment	Test	Assessment
1953 JUNE (Age 8)	KMAA .57 (858)		KGWRV .64 (815)			
			KCA .59 (852)			
	KRAA .63 (840)		KSA .62 (856)		KIA .68 (819)	
1954 JUNE (Age 9)	KMAB .72 (1058)		KGWRB .72 (1061)		KIBV .76 (740)	
			KCB .74 (1053)		KIBNV .68 (736)	
	KRAB .75 (1054)		KSB .68 (1052)		KIB .79 (1042)	
1955 JUNE (Age 10)	KMAA .75 (1097)		KGWRA .73 (1089)		KIAV .75 (906)	
			KCA .76 (1096)		KIANV .69 (905)	
	KRAA .79 (1094)		KSA .70 (1093)		KIA .79 (1077)	
1956 JANUARY (Age 11)	KMAB .77 (1216)	KMAB .74 (1217)	KGWRB .75 (1212)	KGWRB .72 (1215)	KAITV .84 (848)	KAITV .78 (846)
			KCB .84 (1217)	KCB .77 (1220)	KAITNV .74 (847)	KAITNV .68 (845)
	KRAB .80 (1210)	KRAB .77 (1211)	KSB .71 (1216)	KSB .71 (1219)	KAIT .84 (1200)	KAIT .79 (1200)
1956 MARCH KJE (Aged 11+) Selected Group*	Supp. .II .39 (331)	Supp. II .34 (329)	Supp. I .51 (333)	Supp. I .44 (336)	KAIT .61 (536)	KAIT .51 (532)
			Supp. III .27 (331)	Supp. III .36 (329)	Supp. IV .05 (125)	Supp. IV .07 (123)

\* All correlations are uncorrected.

\* All correlations are uncorrected for the effect of selection.



*(b) Correlations.*

The product-moment coefficients of correlation were calculated between each of the M, E and I tests at age 14 and each of the five attainment and the two intelligence tests taken at ages 8, 9, 10 and 11 (English with English, Arithmetic with Arithmetic, etc.).

Generally, the coefficients decreased as the time interval increased from 4 to  $6\frac{1}{2}$  years and there was little change in the order of forecasting efficiency from 8 to 11. In each of these years the highest correlations with the nearest corresponding test at 14 were provided by the tests of intelligence, reasoning arithmetic and comprehension. (For instance, for intelligence, the correlations for intervals of  $6\frac{1}{2}$ ,  $5\frac{1}{2}$ ,  $4\frac{1}{2}$  and 4 years, were respectively : 0.68, 0.79, 0.79, 0.84.)

A comparison was then made of the relative efficiency of each of the 11 year old tests in predicting the tests and Head's assessments at 14. Correlations were always higher for tests at 11 and tests at 14 than for tests at 11 and assessments at 14, but the order of forecasting efficiency of the separate tests at 11 was the same in each case. (For instance, for intelligence : 0.84 for test and 0.79 for Head's assessments.) Tests involving verbal reasoning were generally more reliable over the years than tests of less verbal and more mechanical skills.

Next, an analysis was made of the correlations of individual tests with the composite criterion (M+E+I) tests at 14.

There was no significant difference between the respective correlations for the three best tests of reasoning (Comprehension, Reasoning Arithmetic, Verbal Intelligence) taken at 11 and the composite (M+E+I) test criterion at 14. (The unboosted correlations were 0.83, 0.84, 0.84, respectively.) Even the least prognostic test (Mechanical Arithmetic) showed a correlation of 0.60 with the composite test criterion  $6\frac{1}{2}$  years later when the best in that year (the educational rating) showed a correlation of nearly 0.74.

This analysis was repeated for the correlations with the composite criterion of (M+E+I) Head's assessments at 14. As for the individual items described above, the order of the primary school tests as regards predictive efficiency was very much the same for assessments at 14 as for tests at 14. (Correlations were, however, generally about 0.04/0.05 less.) The test of reasoning arithmetic alone was almost as good a forecaster of either tests or assessments at 14 as the educational rating (the average of the five attainment tests) in each of the four years in the junior school.

The test of non-verbal intelligence was generally the poorest predictor of the two major criteria but combined with the verbal intelligence test, it provided a slightly better forecast than the verbal test alone.

As far as the procedure for secondary school selection in Kent is concerned, the most important items considered are those which make up the combined score as already described. These are : (a) the intelligence test score ; (b) the adjusted mean of attainment test scores ; and (c) the Head's order of recommendation as quantified on the intelligence test. These scores, however, are only available for children in the top 40 per cent. of the age-group. Correlations were calculated between these items and the secondary school criteria.

TABLE 3

COEFFICIENTS OF CORRELATION BETWEEN JUNIOR SCHOOL TESTS AND SECONDARY SCHOOL CRITERIA.

M+E+I TESTS.

M+E+I HEADS' ASSESSMENTS.

(Number of children in the sample given in brackets.)

See Appendix for meaning of abbreviations.

		Average of M+E+I Autumn, 1959, Tests (Age 14)				Average of M+E+I Autumn, 1959, Heads' Assessments (Age 14)			
<i>Junior School Tests:</i>		<i>June</i> 1953	<i>June</i> 1954	<i>June</i> 1955	<i>Jan.</i> 1956	<i>June</i> 1953	<i>June</i> 1954	<i>June</i> 1955	<i>Jan.</i> 1956
English	KGWR—V.B.A.B.	.66 (814)	.72 (1058)	.73 (1087)	.75 (1209)	.62 (811)	.69 (1059)	.70 (1088)	.71 (1213)
	KC—A.B.A.B.	.64 (850)	.78 (1051)	.79 (1094)	.83 (1214)	.60 (848)	.73 (1051)	.75 (1094)	.78 (1218)
	KS—A.B.A.B.	.65 (854)	.71 (1050)	.73 (1091)	.73 (1213)	.62 (852)	.69 (1051)	.71 (1091)	.70 (1217)
Arithmetic	KMA—A.B.A.B.	.60 (858)	.71 (1056)	.74 (1094)	.78 (1212)	.57 (857)	.69 (1057)	.70 (1094)	.74 (1216)
	KRA—A.B.A.B.	.71 (840)	.81 (1052)	.83 (1091)	.84 (1206)	.68 (838)	.77 (1052)	.78 (1091)	.78 (1210)
E.R.	Educational Rating A.B.A.B.	.74 (831)	.81 (1051)	.83 (1089)	.85 (1208)	.68 (829)	.78 (1051)	.79 (1089)	.80 (1212)
Intelligence	Verbal—B.A.B.		.76 (736)	.76 (903)	.84 (846)		.70 (735)	.72 (903)	.77 (845)
	Non-Verbal— B.A.B.		.66 (734)	.69 (902)	.74 (845)		.63 (733)	.66 (902)	.67 (844)
S.I.S.	A.B.A.—KAIT	.67 (818)	.78 (1039)	.80 (1073)	.84 (1195)	.62 (817)	.72 (1039)	.74 (1073)	.78 (1199)
Stage I Selected sample)*	E.R.				.67 (535)				.60 (532)
	S.E.R.				.68 (533)				.61 (531)
	S.I.S.				.60 (534)				.50 (531)
	S.O.R.				.67 (534)				.59 (531)
	Stage I Score				.71 (534)				.62 (531)
	Paper I				.45 (332)				.43 (331)
Stage II (Further selected sample)*	Paper II				.40 (330)				.31 (329)
	Paper III				.29 (330)				.30 (329)
	Supplementary Test Total				.51 (332)				.49 (331)
	Stage II—Total				.56 (332)				.53 (330)
	Mean coefficient of correlation for record card tests:	.67 (7)	.75 (9)	.77 (9)	.80 (9)	.63 (7)	.71 (9)	.73 (9)	.75 (9)
Excess of tests over Heads' Assessments		.04	.04	.04	.05	—	—	—	—

\* All correlations are uncorrected for the effects of selection.



This group is more homogeneous, of course, than the whole sample, so that inevitably, correlations are very much reduced. For instance, the following table (extracted from Table 3), shows this effect operating for two of the variables :

TABLE 4  
CORRELATIONS BETWEEN TEST TOTAL AT 14 AND ASSESSMENTS AT 11.

	Educational Rating at 11	Standardised Intelligence Score at 11
(a) For whole sample (1,200 children)	0.85	0.84
(b) More able part of sample (530 children) .....	0.67	0.60

One feature of the Kent procedure as already described, is the attempt to compensate for the differential school effect and to eliminate competition between primary schools. To this end, educational ratings are scaled to ensure that the adjusted educational ratings have the same average as the standardised intelligence score school by school. Similarly, the primary Head's order of recommendation is quantified school by school on the basis of the intelligence test score. These two adjusted assessments when added to the standardised test score. These two adjusted assessments when added to the standardised intelligence score determine a single combined score which provides a first source of guidance to the local selection panel in assessing the information on the record card. The following table (extracted in part from Table 3) shows the prognostic value of each of these items for the different criteria at age 14 (the approximate number of children involved in each column of correlations is shown in brackets) :

TABLE 5  
CORRELATIONS BETWEEN TESTS AND ASSESSMENTS AT 11 AND AT 14.

	Total of tests at 14 (530)	Total of Head's assessments at 14 (530)	Head's estimates of number of passes (for those children expected to achieve at least - one pass)	
			at 0 level (390)	at A level (175)
Educational Rating .....	0.67	0.60	0.40	0.43
(1) Standardised Educational Rating .....	0.68	0.61	0.42	0.46
(2) Standardised Intelligence Score .....	0.60	0.50	0.36	0.41
(3) Standardised Order of Recommendation .....	0.67	0.59	0.43	0.43
Combined Score : (1) + (2) + (3) .....	0.71	0.62	0.45	0.49

Comparisons may be made in this table vertically between the different criteria at 11 for the effectiveness of selection but horizontally comparison is

difficult because of the differences between the groups of children involved.

It is clear from the above table that the measure of attainment, after adjustment for differential school effects (i.e., the standardised educational rating), is the most useful of the three items which make up the combined score. Furthermore, the best prediction of each of the criteria at age 14 is provided by this combined score which, if known for the whole unselected sample, would yield a correlation of 0.91 with a multiple criterion of secondary school success combining 14+ tests, Head's assessments and forecasts of O level success.

When the analysis was extended to the supplementary test and interview stage, it was found that little additional information of use for prediction was produced.

Calculations were then based on the correlations available to determine the maximum correlation between criteria at 11 and at 14 for different weights. With best weights of 3 : 1 : 2 for SER, SIS and SOR at 11 and best weights at 14 of 5 : 3 : 2 : 1 for tests, Head's assessments, O level forecast and A level forecast the maximum correlation for the selected group of 530 was 0.781 (which corresponds to a correlation of more than 0.91 for the whole age group over the period of nearly 4 years).

In fact, the correlation would not (theoretically) have been much less for a procedure which was considerably streamlined. For instance, it would not be practicable to base selection on the Reasoning Arithmetic test alone, yet if this had been done (without preliminary warning) the correlation with a multiple criterion of  $2 \times (M+E+I)$  tests +  $(M+E+I)$  Head's assessments + O level forecast would have been 0.86, only 0.05 less for the whole age group than the full procedure.

#### V.—CONCLUSIONS.

The follow-up investigation has enabled a comparison to be made over a range of up to  $6\frac{1}{2}$  years between scores on tests of attainment and intelligence, of reasoning ability and of mechanical skills, and of tests timed and untimed.

The most reliable tests are usually but not always the tests of highest predictive value. The most predictive tests are the tests of reasoning. Whether the content is basically literate or numerate makes little difference provided that the test is predominantly in a verbal form. Whether the reasoning test is timed or untimed, seems to have little significant effect on its predictivity.

The adjustment to the attainment score designed to eliminate competition between primary schools does not reduce the predictivity of the attainment score which is as good if not better in this respect than the intelligence score.

A follow-up study can cover the whole range of ability in all types of secondary school by using a multiplicity of criteria which include the Heads' assessments of pupils against the background of the whole age group.

Differences between the sexes in academic achievement follow a consistent pattern. There are really two populations and we are often misled unless we look at the distribution of the whole range of ability for boys and girls separately.

Finally, it is clear that any increase of accuracy of prediction in Kent by modification of procedure can only be minimal. If there is to be an overlap of achievement between different types of secondary schools (and this must inevitably be appreciable), then there must be a corresponding overlap in the functions of the secondary schools in order to produce a balanced community.



## VI.—APPENDIX.

## (a) THE SIX-YEAR FOLLOW-UP STUDY—PROGRAMME OF TESTS.

Random sample : All Kent children born on the fifth or twentieth day of any month of 1945.

Number in the sample : 1,300 boys and girls.

## (i) PRIMARY SCHOOL RECORD CARD ASSESSMENTS.

*June, 1953* (end of first year of junior school course)

<i>Word Recognition.</i>	Vernon's Graded Word Reading Test.	(KGWRV)
<i>Spelling.</i>	Kent Spelling Test (Form A).	(KSA)
<i>Comprehension.</i>	Kent Reading (Comprehension) Test (Form A).	(KCA)
<i>Mechanical Arithmetic.</i>	Kent Mechanical Arithmetic Test (Form A).	(KMAA)
<i>Reasoning Arithmetic.</i>	Kent Reasoning Arithmetic Test (Form A).	(KRAA)
<i>Intelligence.</i>	Kent Intelligence Test (Form A) (Verbal and Non-Verbal)	(KIA(V)) (KIA(NV))

*June, 1954* (2nd Year).

*Word Recognition.* Kent Graded Word Reading Test (Form B). (KGWRB)

Other tests as for 1953 but in Form (B).

*June, 1955* (3rd Year).

All tests as for 1954 in Form (A).

*January, 1956* (4th Year of junior school course).

Tests given as for June, 1954 (Form B), except

*Intelligence.* Kent Annual Intelligence Test (Verbal + Non-Verbal) (KAIT)

NOTE.—By 1962, the testing programme had been reduced as follows :

Junior Years 1, 2, 3 and 4 : KC-, KRA-, KI-, together with  
in Junior Years 2 and 3 only : KS- and KMA-.

The standardised tests remain in the custody of head teachers, three parallel forms being used in annual rotation. A new K.A.I.T., however, is prepared each year and is not available before a prescribed date. Its administration is preceded, at an interval of just under one week, by a 'teaching' intelligence test for which teachers are given explanatory notes. This is followed by a full-length practice test (a previous year's K.A.I.T.) on the day preceding K.A.I.T. Each linked school possesses a copy of the Committee's manual "Notes on the Record Card" (Parts I and II, with Supplements) which contains full details of the construction, use and standardisation of the tests.

## (ii) SECONDARY SCHOOL ASSESSMENTS.

*December, 1959* (in 4th Year of secondary school course).

## (a) Tests.

- (i) Moray House Adult 1 Intelligence Test.
- (ii) Manchester Comprehension Test.
- (iii) Kent Mathematics Test.

## (b) Assessments by Head of Secondary School.

- (i) Attainment in English.
- (ii) Attainment in Mathematics.
- (iii) General overall ability.
- (iv) Remaining subjects.

*(c) Forecasts by Head of Secondary School.*

- (i) Number of O level passes.
- (ii) Number of A level passes.
- (iii) Date of leaving school.
- (iv) Future career.

*(d) Information supplied by Pupil.*

- (i) Two subjects liked.
- (ii) Two subjects disliked.
- (iii) Degree of difficulty of secondary school work.

**(2) THE SELECTION PROCEDURE IN 1963.**

When the January/February testing has been completed, each primary school head teacher enters on a prepared form (S.40(e)) :

- (i) A school Order of Recommendation (O.R.) for all children with a standard intelligence score (S.I.S.) of 105 or above, although provision is made for children with S.I.S. below 105 to be included if specially recommended.
- (ii) The Standard Intelligence Score.
- (iii) The Educational Rating (E.R.), which is the average of the standardised scores in the comprehension and reasoning arithmetic tests.
- (iv) The type of secondary education (grammar, technical or modern) that the head teacher recommends.

Form S.40 (e) is then sent to the Kent Education Committee's Head Office, where the O.R. is quantified to give the S.O.R., by means of the S.I.S. The E.R. is standardised to give the S.E.R., by an adjustment whereby for each school the average of the attainment scores is made the same as the average of the intelligence test scores. The three items S.I.S., S.O.R., and S.E.R., are then aggregated to give the Combined Score.

On the basis of the Combined Score computed at Head Office, children are provisionally assigned to one of the following lists :

- G—admission to a grammar or to a technical high school, according to parental preference.
- T—admission to a technical high school.
- M—admission to a secondary modern school.

Forms S.40(e) thus annotated are then forwarded to local selection panels (consisting of head teachers (one each) from secondary grammar, technical high, secondary modern, and a primary school outside the area) for consideration. The panels place children about whom they feel a need for further information into a borderzone group (List X). The borderzone group is normally restricted to those children whose records are incomplete in some way and may amount in size to 2 to 3 per cent. of the age-group.

While scores are being computed at Head Office, each local selection panel has under consideration the record cards of all the pupils whose names have been entered on Form S.40(e). When the combined scores are received, the panel members compare them with their own provisional assessments and make their decisions. A confidential advance notification is given by panels direct to head teachers of linked schools to provide them with an opportunity for lodging appeals if they consider that any injustice is liable to result from the decisions proposed. The number of places offered by each selection panel is determined by the need as indicated by the achievement of children in the catchment area. The parents of pupils placed on List X are then notified and the pupils concerned are given Supplementary Tests. They may also be interviewed by the local selection panel, if the panel so decides. The supplementary tests consist of a composition (Paper III) and papers in Verbal Comprehension (Paper I) and Mathematics (Paper II). These aim at testing pupils in depth without undue emphasis on speed and do not attempt to



retest abilities which have already been adequately measured in the tests of mechanical arithmetic and spelling recorded on the record card. After careful consideration by the panel of all information then available, pupils are assessed for education in grammar, technical high and modern schools. Panels then notify the head teachers of linked schools who announce the results to the parents by printed letter.

(3) THE SELECTION PROCEDURE IN 1956.

In principle, the procedures in 1956 followed closely that described above for 1963. In 1956, however, the selection panel selected initially a borderzone of 25 per cent. of the age-group to take Supplementary Tests. On these results the panel selected a narrower borderzone of 5 per cent. of the age-group to be interviewed by the panel. The experience of the intervening years suggested that the borderzone could be appreciably reduced without any significant decrease in accuracy.

The combined score (S.I.S., S.O.R. and S.E.R.) was known in 1956 as the Stage I score. The further addition of the Supplementary Test total was known as the Stage II score.

(An experimental Supplementary Test Paper IV (Non-verbal intelligence) was used in 1956 but was discarded in subsequent years.)

*(Manuscript received 18th January, 1963)*

# ROLE-CONFLICT IN ADOLESCENCE

By F. MUSGROVE

*(Department of Education, University of Leeds)*

**SUMMARY.** The role-conflicts of a group of adolescents (and pre-adolescents) were assessed by asking the subjects to indicate how they would behave ideally, how they thought they in fact behaved, and how they thought various adult authorities and their friends expected them to behave. This was done by ranking various areas of conduct as they themselves, and in their view the adults and friends, evaluated them in their behaviour. The extent of disagreement among the ranks was taken as a measure of conflict.

The adolescents were 470 fifteen-year-old grammar and modern school boys and girls in three industrial towns. A last-year junior school group of forty-seven children and sixty-nine late-teenage male technical college students were also investigated. Role-conflict was greatest among the grammar-school boys and technical college students, least among secondary modern school girls and junior school children. Grammar school boys showed more conflict than modern school boys, grammar school girls than modern school girls. There were no social-class differences within a particular type of school. Teachers, mothers, fathers (and bosses) were seen by all groups to have substantially the same expectations; but there was greater conflict between self-conception and perceived expectations of friends. The grammar school pupils' conception of their role stood midway between the perceived expectations of friends and of adults; the modern school children did not so generally see themselves between such extreme and conflicting demands.

## I.—INTRODUCTION.

THE concept of social role which has been elaborated by American sociologists and psychologists such as Merton (1957) and Levinson (1959) provides a valuable frame of reference for the interpretation of conflict situations. Social institutions confront their members with adaptive dilemmas, and these can be described in terms of the 'role-set' (Merton) and 'role-demand,' 'role-conception' and 'role-performance' (Levinson). The role-demands and expectations made upon a person occupying a particular status (as wife, foreman, teacher, etc.) may not coincide with the person's own conception of his role; and neither his conception nor the expectations of others may match his actual role performance. The size of the discrepancy between role-demands, role-conception, and role-performance is a measure of the conflict experienced by a person in a particular status. But this theoretical model is more complicated than a simple triangle of forces; for the role-demands are not necessarily single or unified: the role-set usually implies an array of possibly conflicting expectations. The individual may perceive not only that he cannot perform as he would wish, but that many people in significant social relationships with him expect him to perform differently.

No work is known to the author which attempts to quantify and measure the conflict or tension between role-demands, role-conception, and role-performance for individuals occupying particular statuses. This paper reports an attempt to measure the conflicts of people occupying the status of adolescent pupil. The conception which schoolchildren and technical college students have of their role was compared with the demands which they perceived various adult authorities and their friends to make upon them, on the one hand, and with their perception of their own role performance, on the other.



## II.—OUTLINE OF PREVIOUS RESEARCH.

There is no dearth of impressionistic accounts of conflict between adolescents and their parents and their peers. Margaret Mead (1948) and David Riesman (1950) have no doubt of their general and early capitulation to the expectations of their friends. Some doubt has been thrown on this picture by Morris (1955, 1958) in England, and more recently by Riley, Riley and Moore (1961) in their attempts to provide empirical verification for Riesman's hypotheses in America. Morris's inquiries into what adolescents thought they would or should do when there were conflicting demands from parents and peers led him seriously to question the potency of peer-group attraction and authority, and the tendency to adolescent conformity, in England; but he found rather more inclination to conform to friends among secondary modern than among grammar school pupils. Riley, Riley and Moore, attempted to establish to what extent 2,500 middle-class American high school students were 'other-' or 'inner-directed' by asking for their evaluation of twenty personality models or 'vignettes.' The vignettes represented 'Achievement Models,' 'Good Time Models,' and 'Peer Relations Models.' The subjects were asked to indicate those they themselves would like to resemble ('self-expectation'), those they thought their popular coevals preferred as friends ('perceived peer expectations'), those they thought their parents wanted them to be like ('perceived parent expectations'), and those they thought it would be helpful to be like later on, after school ('adult self-expectation'). There was a tendency for these middle-class American adolescents to stand midway between the perceived expectations of friends and of parents: (for 67 per cent. to prefer the 'success model,' for example, while only 48 per cent of them thought that their friends wished them to resemble this model, but 80 per. cent. thought that their parents wished them to do so). There was a sense in which 'the adolescent's own values seem to form a bridge between the perceived other-direction of his peers and the perceived inner-and-other-direction of his parents.'

## III.—PLAN OF THE INVESTIGATION.

The inquiry reported below was carried out by the author in the grammar school and a secondary modern school in each of three widely separated industrial towns in the North and Midlands. Each town is engaged in mining and/or heavy industry with a population of approximately 40,000 inhabitants. Each town has a medium-sized mixed grammar school recruited through orthodox selection procedures at 11+. The secondary modern schools which co-operated in the research were also mixed. In addition, one junior school and one college of technology were included in the inquiry.

A questionnaire was designed to make possible the measurement of role-conflict among the pupils in these eight educational institutions. The definition of adolescent role was derived from a previous inquiry by the author (Musgrove, 1963) into the self-pictures of adolescents. In this earlier inquiry young people between the age of 12 and 16 had been asked to complete in any way they wished, which gave a true impression of their views, three sentences beginning: "Boys (or girls) of my age are . . ." "Boys (or girls) of my age can . . ." "Boys (or girls) of my age should . . ." Four areas of behaviour were mentioned with the



greatest frequency in completing the third sentence ; these were made the basis of the questionnaire :

- (1) " Fit, and good at games."
- (2) " Behave sensibly and generally ' act your age '."
- (3) " Be respectful, courteous and polite to your parents and elders."
- (4) " Be quite free to stay up late, or to stay out late, if you wish."

These four statements were taken to be descriptive of a generalized adolescent role-conception.

The subjects in the present inquiry were asked to rank these four areas of behaviour to show what they ideally wished to be like (' role-conception '), what they thought mothers, fathers, teachers, friends (and bosses in the case of the part-time technical college students) respectively wished them to be like (' role-expectations '), and what they thought they were in fact like (' perceived role-performance '). Thus, each subject was asked to place in the first column (Ideal Self) " 1 against the statement you would most like to apply to you, 4 against the statement which you are least bothered to be, 2 against the statement you would like to apply to you after number 1, and 3 against the statement you would like to apply to you after that." Similar detailed instructions were given for the columns for ' Mother,' ' Father,' ' Teachers,' ' Boss,' and ' Friends,' e.g., " In column 2 show what you think your father wants you to be like : put 1 against the statement which you think he would most like you to be, 4 against what you think he is least bothered about, 2 against what you think he wants you to be after number 1, and 3, after that." For the column relating to ' Actual Self ' the instructions were : " In column 6 put 1 against the statement which you think as a matter of fact fits you best, put 4 against the statement which you think as a matter of fact fits you least, 2 and 3 against the statements which fit you second and third."

By computing the coefficient of concordance (Kendall, 1948) for the six (or seven) ranks, a measure could be obtained for the degree of an individual's role-conflict ; by computing the coefficient for the perceived expectations of all three (or four) adults, a measure could be obtained of the felt congruence of demands from adult authorities. Differences in measured conflict could be established between age-groups, sex-groups, and between adolescents of the same sex and age in different types of educational institution. The occupations of subjects' fathers were also obtained and these were graded on the Registrar-General's scale of occupations, Class III being sub-divided into IIIA, routine non-manual, and IIIB, skilled manual. It would thus be possible to establish the incidence of conflict within different ' social class ' groups in different types of school.

#### IV.—STATEMENT OF RESULTS.

The questionnaire was completed by the children in the fourth year of the grammar schools (N=275) and in the fourth year of the modern schools (N=194) by the forty-seven children in the last year of the junior school, and by sixty-nine male students aged 17 to 20 following part-time courses for professional qualifications in engineering in the technical college. The average age of the secondary school boys and girls was 15 years, of the junior school children 11 years.

##### (a) *Differences in Overall Conflict.*

The greatest degree of conflict was found among grammar school boys and the (male) technical college students, the least among secondary modern girls



and junior school children. Grammar school girls and secondary modern school boys occupied an intermediate position. Comparisons were made by calculating the coefficient of concordance,  $W$ , for each pupil in a random sample of 1 in 2 from each sex-group in every institution. Although there were highly significant differences between types of school, there were no differences between the grammar school children in the three towns, or between the secondary modern children. The results for grammar schools are, therefore, combined, and for modern schools.

TABLE 1  
PERCENTAGE OF PUPILS AT THREE LEVELS OF CONFLICT.

Value of $W$	Gr. Boys (N 64)	Mod. Boys (N 50)	Gr. Girls (N 78)	Mod. Girls (N 49)	Tech. Boys (N 35)	Jun. Boys (N 12)	Jun. Girls (N 12)
High Conflict 0.00—0.33 . . . . .	56.3	32.0	39.7	12.2	57.1	16.7	8.3
Medium Conflict 0.34—0.66 . . . . .	39.1	44.0	35.9	20.4	34.3	16.7	41.7
Low Conflict 0.67—1.00 . . . . .	4.6	24.0	24.4	67.4	8.6	66.6	50.0

More grammar school than secondary modern boys showed high levels of role-conflict ( $\chi^2=26.0$ , d.f.2,  $P<0.001$ ), more grammar school girls than secondary modern girls ( $\chi^2=23.33$ , d.f.2,  $P<0.001$ ). The technical college students showed the same high levels of conflict as the grammar school boys, and the latter greater conflict than the grammar school girls ( $\chi^2=10.80$ , d.f.2,  $P<0.01$ ). The grammar school girls did not differ significantly from the modern school boys, but the latter showed greater conflict than modern school girls ( $\chi^2=18.7$ , d.f.2,  $P<0.001$ ). Junior boys showed less conflict than modern school boys ( $\chi^2=8.95$ , d.f.2,  $P<0.05$ ), but neither junior school boys nor junior school girls differed significantly from the low level of conflict found in the modern school girls. When  $W$  was calculated from the median ranks for both sex groups in each type of school, the following coefficients were obtained: grammar school boys 0.25 (N.S.), technical college students 0.32 (N.S.), grammar school girls 0.37 (N.S.), secondary modern boys 0.33 (N.S.), secondary modern girls 0.63 (significant at the 0.01 level:  $F=8.5$ , d.f.2.66/13.30,  $F_{0.01}=6.0$ ), junior school boys 0.66 (significant at the 0.01 level:  $F=9.3$ , d.f.2.66/13.30,  $F_{0.01}=6.0$ ), junior school girls 0.96 (significant at the 0.01 level:  $F=95$ , d.f.2.66/13.30,  $F_{0.01}=6.0$ ).

Oppenheim (1958) has suggested that working-class children in grammar schools and middle-class children in modern schools may experience conflicts of value which do not beset middle-class children in the former type of school or working-class children in the latter. No evidence was found in this inquiry to support this view. Thus, in one grammar school the coefficient of concordance calculated from the median rankings made by the twenty-two middle-class boys (whose fathers were placed in the Registrar-General's occupational grades IIIA-I) was 0.12; but for all the boys in the school ( $N=39$ ), the value was 0.10. The coefficient for the twenty-eight middle-class girls in another grammar school was 0.17, for all fifty-eight girls in the school, 0.22. The value of  $W$  for twelve boys from white-collar homes in one modern school was 0.88, for all

thirty-six boys in the school, 0.69; the value of *W* for ten 'white-collar' girls in a modern school was 0.60, for all thirty-two girls in the school, 0.81.

(b) *Specific Areas of Conflict.*

No group perceived marked differences in the demands of the various adult figures, fathers, mothers, teachers, and, where applicable, bosses. "Behave sensibly and 'generally act your age'" and "Be respectful, courteous and polite to your parents and elders," were seen as the prime expectations of all adults: the median rank accorded by all groups of subjects was either 1 or 2. All groups saw all adults as placing "Fit and good at games" third in importance, and being "Quite free to stay up late, or to stay out late, if you wish" fourth. On the other hand, there was a high degree of conflict between the subjects' perceived expectations of their friends, their own preferred behaviour, and their conception of their actual behaviour. This conflict was much greater among grammar school (and technical college) students than among modern school pupils and centred particularly on being "courteous, etc., to parents and elders" and being "free to stay up, or stay out, late." The grammar school boys saw "being free to stay up, or to stay out, late" as an important expectation of their friends (median rank 2), ideally they attached little importance to this in their own behaviour (median rank 4), but they perceived their actual behaviour midway between their own ideal and the expectation of their friends (median rank 3). There was a similar conflict over being polite and respectful to parents and elders: ideally they rated this high, saw their friends as rating it low, and gave it an intermediate ranking as a feature of their actual behaviour. Secondary modern girls (and to a smaller extent secondary modern boys) showed less conflict over both areas of behaviour: they did not attach much importance to staying out late, they did not rate this as an outstanding feature of their actual behaviour, and did not see it as a prime demand by their friends. They saw their friends attaching the same high importance to respect and politeness to parents that they did themselves.

TABLE 2  
CONFLICT BETWEEN IDEAL SELF, ACTUAL SELF, AND FRIENDS.

Subjects	'Be respectful, courteous and polite, etc.' (Median Ranks)		
	Ideal Self	Actual Self	Friends
Gr. Boys (N 123)	2	3	4
Gr. Girls (N 152)	2	2	3
Mod. Boys (N 99)	2	2	3
Mod. Girls (N 95)	2	2	2
	'Be quite free to stay out late, etc.'		
	Ideal Self	Actual Self	Friends
Gr. Boys	4	3	2
Gr. Girls	3	3	2
Mod. Boys	4	3	2
Mod. Girls	4	4	3



The median rankings of adults' expectations over the four areas of behaviour were in close agreement, the median rankings for ideal self, actual self, and friends, showed a high degree of conflict :

TABLE 3

AGREEMENT BETWEEN ADULT EXPECTATIONS, CONFLICT BETWEEN SELF AND FRIENDS.

	Gr. Boys	Tech. Boys	Gr. Girls	Mod. Boys	Mod. Girls
$W_A$ .....	0.80	0.90	0.93	0.85	0.85
$W_{SI, SA, F}$ .....	0.18	0.07	0.18	0.09	0.53

Legend :

A—Adults.

SI—Self, Ideal.

SA—Self, Actual.

F—Friends.

(c) *Differences in Expectation between Grammar and Modern Pupils.*

Sex groups in different types of school differed in the conflict in their perceived demands of adults, friends, and their ideal selves ; but even when the expectations did not conflict, they differed in the weight which they saw friends, parents, teachers and themselves attaching to the four areas of behaviour. Grammar school boys themselves attached more weight to being fit and good at games than did modern school boys ; and more than the modern school boys they saw their friends as making demands in this field of behaviour. Grammar school boys saw adults as having a greater expectation that they should 'act their age,' and friends a smaller expectation that they should be courteous and polite to their elders and parents. Modern school boys attached more importance to respect and politeness in their own ideal behaviour, and saw their friends attaching less importance to their freedom to stay up or to stay out late.

Like the grammar school boys, grammar school girls attached greater weight than the modern school girls to being fit and good at games, and saw their friends as having a higher expectation in this regard ; they saw their friends as attaching less importance to 'acting your age' and to being polite and showing respect. They attached more importance to being free to stay up, or stay out, late, and saw their friends as placing more weight on their being able to do so.

Like the middle-class adolescents in the American inquiry reported by Riley, Riley and Moore, both grammar school boys and girls constituted, in their ideal conception of themselves, a bridge between the widely different demands they perceived from adults on the one hand, and from friends on the other. This was not so generally the case with secondary modern boys, and still less so with secondary modern girls.

Table 5 shows not only where 'self-expectation' is located in relation to the perceived expectations of adults and of friends, but the 'spread of conflict'—particularly wide with grammar school boys, particularly narrow with secondary modern school girls. The latter are much nearer to the expectations of adults than of friends ; in general the former are nearer in their preferred behaviour to the perceived expectations of friends than of adults. Morris found rather more 'conformity to friends' among secondary modern than among grammar school pupils ; it is true that the present inquiry shows that grammar school

TABLE 4

SIGNIFICANT DIFFERENCES BETWEEN GRAMMAR AND MODERN PUPILS IN PERCEIVED EXPECTATIONS.

(Percentages ranking an area of behaviour high : 1 or 2.)

Area of Behaviour	Boys			
	Grammar (N 123)	Modern (N 99)	$\chi^2$	P
'Fit and good at games':				
Ideal Self .....	60.8	47.5	6.10	0.05
Friends .....	80.5	55.5	16.20	0.001
'Sensible and "act your age"':				
Adults .....	88.1	74.1	18.85	0.001
Friends .....	28.4	51.5	11.50	0.001
'Respectful, polite, etc.':				
Ideal Self .....	56.1	76.7	10.35	0.01
Friends .....	12.2	37.3	18.93	0.001
'Free to stay up or stay out':				
Friends .....	94.3	57.3	11.34	0.001
	(N 152)		GIRLS (N 95)	
'Fit and good at games':				
Ideal Self .....	27.0	9.5	8.34	0.01
Friends .....	40.1	25.2	5.57	0.05
'Sensible and "act your age"':				
Friends .....	57.9	77.0	9.25	0.01
'Respectful, polite, etc.':				
Friends .....	31.5	67.3	29.50	0.001
'Free to stay up or stay out':				
Ideal Self .....	38.8	9.5	16.86	0.001
Friends .....	65.8	29.5	30.80	0.001

TABLE 5

SELF-CONCEPTION IN RELATION TO PERCEIVED EXPECTATIONS OF ADULTS AND OF FRIENDS.

Percentages ranking an area high : (1 or 2.)

School-Sex Group	I.—Fit, etc.			II.—Sensible			III.—Polite			IV.—Stay Out		
	A	S	F	A	S	F	A	S	F	A	S	F
Gr. (B.) N=123. ....	19.2	60.2	80.5	88.1	53.7	28.4	90.2	56.1	12.2	3.0	30.1	94.3
Mod. (B.) N=99 ..	27.6	47.5	55.5	74.1	50.5	51.5	84.8	76.7	37.3	8.1	26.2	57.5
Gr. (G.) N=152. ....	14.0	27.0	40.1	94.3	67.1	57.9	96.0	73.7	31.5	2.0	38.8	65.8
Mod. (G.) N=95 ..	15.4	9.5	25.2	87.0	90.5	77.0	91.2	87.3	67.3	4.0	9.5	29.5

Legend :

A—Perceived expectations of adults.

S—Self-expectation.

F—Perceived expectations of friends.



boys (and to a lesser extent grammar school girls) are 'further from' their friends than their counterparts in the modern schools. But they are also much further from adults. In particular, the grammar school boy's role-conception is remote from the perceived role-expectations of adults and friends: he is located at the point of maximum tension. At the other extreme the secondary modern girl's role-conception is close to the perceived role-expectations of both adults and friends. Thus, while 19.2 per cent. of grammar school boys thought that adults had a high expectation of their being fit and good at games, and 80.5 per cent. thought their friends had this high expectation, an intermediate proportion, 60.2 per cent. rated this as important in their ideal behaviour. But modern school girls were not a 'bridge,' in this sense, between adults and friends: while 9.5 per cent. attached importance to being fit and good at games, 15.4 per cent. thought that adults had a high expectation of this, and 25.2 per cent. thought that their friends had. Whereas, the spread of conflict for the modern school girls (the difference between highest and lowest perceived expectations) was only 15.7 per cent. for the grammar school boys the spread of conflict over this area of behaviour was 61.3 per cent.

#### V.—DISCUSSION AND CONCLUSIONS.

The interpretation of the findings reported in this paper can be only speculative. The role-conflicts of boys were consistently very high in three widely separated grammar schools; the role-conflicts of girls in three modern schools in the same areas were consistently very low. Intermediate in degree were the role-conflicts of grammar school girls and secondary modern boys. Grammar school boys and girls were significantly more at odds with themselves, with their friends, and with teachers, mothers and fathers, than were their counterparts in the modern schools.

The work of Hallworth (1961) might lead us to expect that the grammar schools select personalities already prone to self-blame, self-criticism, and a hypercritical attitude to others. But if this were the explanation of these findings, we might reasonably expect that those children in the junior school who had been assessed to have the intelligence and academic aptitude for a grammar school education would already experience more role-conflict than their unselected contemporaries. This was not the case. Two boys and five girls in the junior school had been selected for grammar school places: they were from various social levels, their fathers being an electrician, a van driver, a greengrocer, a skilled factory operative, a baker, a hosiery knitter, and a fitter, respectively. The coefficient of concordance for each child was as high as for their classmates: 0.83, 0.83, 0.61, 0.83, 0.88, and 0.95.

It seems probable that it is the experience of a particular type of educational institution, rather than pre-existing personality traits, or social background, which promotes or reduces role-conflict. The grammar school and the technical college, which make extreme demands upon their pupils and emphasize their dependence and protracted exclusion from full involvement in adult affairs, may induce a deeper sense of conflict than the modern school, with its more moderate demands and more immediate relationship, particularly for the 14 and 15-year-olds, with the adult world. The male technical college students were older than the grammar and modern school pupils and came from both modern and grammar schools; yet, irrespective of their previous education, their role-conflicts were as extreme as those found among the grammar school boys.

Bene (1957) found that grammar school boys had less positive feelings towards school than modern school boys and perceived negative feelings to come from adult authorities more often; they were more critical of the manners, behaviour and habits of children and people generally, and, indeed, more negative in their feelings towards the world as a whole. It was the grammar school boys who, in general, felt most resentment towards their environment. This was not explicable in terms of social background: there were no significant differences between middle-class and working class boys. Bene's findings seem to be congruent with, and to lend some support to, the data reported in this paper. (Her investigation concerned boys only.) Although the grammar schools carry higher prestige than the modern schools and promise better opportunities in the future, although their pupils are the 'successful' and those of the modern school the 'rejected,' it is likely that the social and educational experiences they provide within their walls promote anxieties, uncertainties, tensions and conflicts. The processes and circumstances which are responsible are an urgent matter for research.

#### VI.—REFERENCES.

- BENE, E. (1957). The objective use of a projective technique, illustrated by a study of the differences in attitudes between pupils of grammar schools and secondary modern schools. *Brit. J. Educ. Psychol.*, 27, 89-100.
- HALLWORTH, H. J. (1961). Anxiety in secondary modern and grammar school pupils. *Brit. J. Educ. Psychol.*, 31, 281-291.
- KENDALL, M. G. (1948). *Rank Correlation Methods*. London: Griffin.
- LEVINSON, D. J. (1959). Role, personality, and social structure in the organizational setting. *J. Abnorm. Soc. Psychol.*, 58, 170-180.
- MEAD, M. (1948). Social change and cultural surrogates. In *Personality in Nature, Society and Culture*, ed. Kluckhohn, C., and Murray, H.A. New York: Knopf.
- MERTON, R. K. (1957). Role-set: problems in sociological theory. *Brit. J. Soc.*, 8, 106-120.
- MORRIS, J. F. (1955). *A Study of Value-judgments in Adolescents*. Ph.D. thesis (unpublished), University of London.
- MORRIS, J. F. (1958). Symposium: the development of moral values in children. II.—The development of adolescent value-judgments. *Brit. J. Educ. Psychol.*, 28, 1-14.
- MUSGROVE, F. (forthcoming 1963). Inter-generation attitudes. *Brit. J. Soc. Clin. Psychol.*
- OPPENHEIM, A. N. (1958). Social status and clique formation among grammar school boys. *Brit. J. Soc.*, 4, 228-245.
- RIESMAN, D. (1950). *The Lonely Crowd*. New Haven: Yale University Press.
- RILEY, M. W., RILEY, J. W., and MOORE, M. E. (1961). Adolescent values and the Riesman typology: an empirical analysis. In *Culture and Social Character*, ed. Lipset, S. M., and Lowenthal, L. Glencoe, Ill.: The Free Press.

(Manuscript received 25th January, 1963.)



# A COMPARISON OF PROGRAMMED LEARNING AND NOTE-MAKING AT TWO AGE LEVELS\*

By PAMELA K. POPPLETON AND K. AUSTWICK  
(*Department of Education, University of Sheffield*)

**SUMMARY.** Two investigations were undertaken to examine learning from two different forms of programmed instruction. In one condition, subjects were asked to make continuous responses to a linear programme; in the other, subjects were asked to work through the same programme presented in the form of a text-book, and to make their own notes.

Differences in learning were examined at two age levels. Twenty post-graduate students in a Department of Education took part in one investigation, and thirty-six first year grammar school pupils formed the subjects for the other.

The students reached the same level of achievement from both forms of the programme, but the grammar school pupils who worked from the programme learned better than those who worked from the text.

Student comments on learning by these methods are described and the use of the efficiency index as an evaluative measure is discussed.

## I.—INTRODUCTION.

THE response requirement in programmed instruction is thought to be basic to efficient learning. The learner is asked to respond overtly to each item, either by writing in some missing words, or by pressing a button and, in this way, it is assumed that he is active and paying attention to the material in front of him. All techniques enable him to compare his answer with the correct one immediately and he thus has maximum guidance and reinforcement from the continuous knowledge of his own progress. Holland (1960) found that this kind of response requirement was necessary, provided that it ensured careful reading of the items.

Subsequent studies have thrown some doubt upon the effectiveness of the overt response, particularly in relation to constructed response programmes of the kind advocated by Skinner (1958). Investigators have compared the performance of groups of learners who have made continuous overt responses to each frame with that of groups who have simply read through the programme and mentally filled in the blanks (Evans, Glaser and Homme, 1960; Silverman and Alter, 1961; Roe, *et al*, 1960). A variety of subjects have been covered by programmes in these investigations and a general summary of the results indicates that the covert, or no-response condition achieved equally efficient learning with a considerable saving of time, this being particularly so with easy subject matter (Goldbeck, 1960).

A more realistic condition of working in the ordinary course of study is for students to make notes as they read in order to summarise the important points. It is assumed that in doing so they learn more efficiently and provide themselves with a framework for future revision. In these circumstances, learners are making overt responses which they can check from the text, but the responses are not guided systematically.

\* These investigations were made possible by a grant from the Research Fund of the University of Sheffield.

The making of notes from books is itself a learned skill which is usually picked up in a haphazard way and which becomes highly idiosyncratic. Very little is known about its efficiency. It is laborious and the motivation is entirely dependent upon the learner rather than upon any manipulation of the learning situation, though it is reasonable to suppose that it will be easier the more the material has been structured by the writer. It is also likely to be a more successful technique in the case of mature students who may be assumed to be skilled in learning from the written word than in the case of younger learners. It was, therefore, decided to undertake two investigations, one with university students, and one with grammar school pupils.

The purpose of the investigations was to compare learning from a text with an overt response in the form of note-making with learning from a constructed response programme in which the responses are determined by the programme itself. In order to do this, it was necessary that both programme and text should cover exactly the same topics and so the texts used consisted of the programmes which had the blanks filled in. The result was a book which differed from an ordinary text book only by virtue of the fact that it was sequenced, i.e., the units of information were assembled in a pre-arranged sequence with a minimum of cross reference. This is no doubt an ideal towards which textbook writers strive, and it is difficult to assess to what extent such a text would depart from an ordinary textbook on the subject. It is, however, the only kind of text with which valid comparisons can be made.

## II.—PLAN OF THE INVESTIGATIONS.

### (i) *Subjects and Materials.*

Students reading for the post-graduate diploma in Education at the University of Sheffield acted as subjects in the first investigation. During the diploma year students take optional courses among which was a course on Experimental Investigations and one on Teaching Machines. Twenty male students who enrolled for both these courses took part in the experiment. A record of their proficiency in mathematics and statistics was available.

The number of students who participated was necessarily limited to those enrolling in these two courses and limited by of the difficulty in obtaining sufficient copies of a tested programme. In this investigation a constructed response programme in Elementary Statistics was used as it would form an integral part of the normal work of these groups. It was devised by Dr. J. Annett of the Department of Psychology and had been pretested and revised. It was prepared in two forms: as originally designed with a response requirement at each step (Form A), and as a text with the missing words filled in (Form B). Both forms were presented as booklets which were as identical in layout as possible.

Subjects in the second investigation were forty first-year pupils at a maintained co-educational grammar school. The subjects were drawn from unstreamed classes and all had followed the same course in mathematics. They covered the range of mathematical achievement as shown by an examination taken at the end of the term prior to the investigation.

The programme used in this case was a constructed response programme on the multiplication of bracketed terms devised by one of the writers. This was also pretested and revised and was prepared in two forms as for Experiment 1.



(ii) *Experimental Procedures.*

As the experimental groups were small, a matching procedure was adopted. In both experiments, a short test was given before the experiment which covered the subject matter of the programme, and it was found possible to match pairs of subjects on the result in Experiment 1, taking into consideration also their level of achievement in mathematics. In Experiment 2, the pretest simply showed that all subjects were ignorant of the content of the programme and their scores were zero. In this case, therefore, matching was accomplished on the basis of the previous mathematics examination. Each member of a pair was randomly allocated to one of the two treatment groups. It was thought possible that at this age, success on programmed material might be related to verbal comprehension. The subjects were, therefore, given the Watts-Vernon Reading Test (1950), and the means and standard deviations for the two groups were practically identical.

In each investigation, the control or 'programme' group was asked to work through the programme in form A, writing in responses, while the members of the experimental or 'text' group were told to use the book as they would use any textbook and to make any notes that they wished. Instructions were standardised and an example from form B is given below:

"You are going to learn how to multiply brackets in algebra.

Your lessons are inside this folder.

They are divided into small sections called *frames*.

We want you to read through each frame carefully, making sure that you understand the working.

As you do this you can make *any notes you like*.

To make a note you can copy out any of the examples, or make up your own, or write out the words which you want to remember."

Some examples were then given for guidance.

Both groups worked under exactly the same conditions. The programmes, texts and notes were taken away at the end of each session, and in Experiment 1 students were instructed not to consult any other sources between sessions. All relevant textbooks were removed from the library and students appeared to be delighted to comply with this novel request!

The first experiment continued over five hourly sessions of one per week, between October 5th and November 2nd, 1961. One or two students put in extra time in order to complete the programme. The second experiment also occupied five sessions, but in this case each session consisted of the normal lesson time of forty minutes and it was necessary to concentrate all five into one week of the Summer term, 1962. Again, by making more time available to one or two individuals, all were able to complete the programme.

Because of different completion times for subjects, it was not possible to give a test immediately after the work was finished and a day or two was allowed to elapse before the post test was given. In order to examine the retention of what had been learned, the test was administered again after a period of time during which no work was done on the topics. In Experiment 1, ten weeks intervened before the groups met again and were tested. In Experiment 2, however, the maximum period which could be allowed was six weeks and the test was administered just before the Summer holiday.

The tests used in both investigations were designed to sample the whole of the material covered. Fry (1960) has shown that the type of test item used is likely to affect the result, and so three types of item were included. Thirty items were divided into three groups: ten constructed response items taken from the programme, ten multiple choice items and ten problems which required the application of what had been learned. Thus, the tests were, as a whole, aimed at evaluating subjects' ability to recall, recognise and transfer their knowledge.

## III.—STATEMENT OF RESULTS.

TABLE 1  
EXPERIMENT 1

MEANS AND STANDARD DEVIATIONS OF RAW SCORES, TIME TAKEN TO WORK THROUGH THE PROGRAMME, AND EFFICIENCY INDICES CALCULATED ON THE POST TEST SCORES.

		Pre test	Post test	Retest	Time in minutes	Efficiency Index
Programme	M. ....	3.90	29.10	26.65	262	8.56
	S.D. ....	2.90	6.56	6.63	54.26	3.78
	N. ....	10	10	8	10	10
Text	M. ....	3.40	32.80	30.00	248	9.94
	S.D. ....	3.04	3.58	4.47	64.94	3.05
	N. ....	10	10	8	10	10
Wilcoxon test	T	—	14	11	13	14
	No. of pairs	—	10 N.S.	8 N.S.	10 N.S.	10 N.S.

TABLE 2  
EXPERIMENT 2

MEANS AND STANDARD DEVIATIONS OF RAW SCORES, TIME TAKEN TO WORK THROUGH THE PROGRAMME AND EFFICIENCY INDICES CALCULATED ON THE POST TEST SCORES.

		Maths. Exam.	Vernon-Watts Reading Test	Post test	Retest	Time in minutes	Efficiency Index
Programme	M. ....	56.6	25.5	17.41	15.29	175.82	10.20
	S.D. ....	15.00	2.95	8.94	7.94	16.49	6.68
	N. ....	20	20	17	17	17	17
Text	M. ....	56.7	24.9	11.42	11.68	155.42	7.58
	S.D. ....	15.28	2.92	6.59	8.38	18.93	4.69
	N. ....	20	20	19	19	19	19
Wilcoxon matched pairs signed-ranks test	T	—	—	13†	35	26*	26*
	No. of pairs	—	—	16	16 N.S.	16	16

\* Significant at  $p = .025$  (one-tailed).

† Significant at  $p = .005$  (one-tailed).



(a) *Comparison between treatment groups on the post test.*

Because the groups taking part in these investigations were small, it was felt that the data would be analysed most appropriately by non-parametric statistics. Accordingly, the Wilcoxon signed-ranks test for matched pairs (Siegel, 1956) was used to compare the performance of the treatment groups. The 'No. of pairs' in all Tables refers to those pairs for which differences existed.

The efficiency index shown in these tables is a measure which has been used by American investigators in an attempt to combine the criteria of test score and learning time. It is calculated from the formula:

$$\frac{\text{Test score}}{\text{Time taken on programme}} \times 100$$

In Experiment 1, no differences emerge between the treatment groups on any of the three measures. In Experiment 2, the programme group scored significantly higher than the text group on the post test, but took significantly longer to work through the material. Nevertheless, when these measures are combined in the efficiency index and groups compared again, the programme group maintains its superiority.

(b) *Comparison between groups on the retest.*

The results of the retest given to examine retention of the material, shows that in neither case were differences to be found between the treatment groups after a period of time.

In Experiment 1, the mean score of both groups had dropped and this fall was found to be significant at  $p < .01$  so that all subjects had forgotten a certain amount. In Experiment 2, however, neither group showed a significant loss, though the mean score of the programme group was lower than on the post test, whereas that of the text group increased slightly. On the retention test, therefore, the previously observed differences between the treatment groups disappeared.

(c) *The effect of type of test item.*

TABLE 3  
MEAN SCORES ON THREE TYPES OF TEST ITEMS.

	Experiment 1			Experiment 2		
	Con- structed Response	Multiple Choice	Problems	Con- structed Response	Multiple Choice	Problems
Programme N . . . . .	7.40 10	6.76 10	6.07 10	5.65 17	5.88 17	5.90 17
Text N . . . . .	8.25 10	7.12 10	7.57 10	4.58 19	3.89 19	2.98 19
Wilcoxon T . . . . .	8.5	22.5	19	17*	15†	9.5†
test (between groups). No. of pairs..	8	10	10	13	14	16
K-W (Analysis of variance between items).	—	10.3†	—	—	2.16	—

\* Significant at  $p = .05$ .

† Significant at  $p = .01$

The null hypothesis to be tested was that the type of test item used was not important in establishing the criterion of learning. The responses to the three types of item were, therefore, compared by the Kruskal-Wallace one-way analysis of variance for ranked data (Siegel, 1956). A significant difference was found in the case of Experiment 1, but not in Experiment 2. The source of the difference in the first investigation was found to be superior performance on constructed response items as against problems. When treatment groups were compared on each item, no differences were apparent. In the second investigation, however, the programme group did significantly better than the text group on both multiple choice items and problems and showed a more consistent level of performance throughout.

#### IV.—DISCUSSION AND CONCLUSIONS.

Whilst every attempt was made to ensure that these two investigations were parallel in all respects, it was found that the demands of school and university organisation made this extremely difficult. We set out to examine the effect of two methods of working at two different age levels. Although the results obtained from small samples cannot be more than suggestive, at the same time, they are in line with what might be expected in the light of everyday experience.

It appears that both methods of working were equally efficacious with the group of post-graduate students. Where note-taking skills are sufficiently developed, no particular advantage is gained by requiring a response at each step of the programme. Students were invited to comment on both forms of the programme. Some students in each group expressed satisfaction with the course, and in the programme group the consensus of opinion was that it would be useful as a revision exercise and would also aid understanding of lectures and textbooks. One or two students said that after the novelty wore off, it became a little boring and tended to become a mechanical exercise which involved little thought. The students in the text group thought that, as a text, there was too much repetition. Generally, they felt that they had achieved good understanding of the topic but doubted whether they had really learned it. The notes made by these students were examined and it was found that they had made notes on approximately 43 per cent. of the items. In view of the fact that they had learned the topic as well as the programme group, the claim of excessive repetition was probably justified. It also suggests that the original programme could be revised successfully with larger steps. There was not a great deal of variation among individuals in the amount of note-taking. The fewest items to be noted out of a total of 206, was 66 and the greatest, 106. A rank order correlation between the number of items noted and success on the post test was 0.60 ( $p < .01$ ). Within this fairly narrow range, those who made the most notes did best.

In the case of young grammar school pupils, a different picture emerges. In Experiment 2 the programme group was more successful than the text group but took longer. Comments from the pupils were difficult to evaluate, as they all politely said that they had enjoyed the work and they had probably welcomed a change from the usual classroom routine. All worked willingly throughout the course of the experiment.

The notes made by the text group showed the undeveloped skill that one would expect at this stage. The first few sessions were occupied by copying down everything in the text, though later on most pupils had become more selective and speeded up their working considerably. There was great individual



variation. The mean number of items noted was sixty from out of a total of 119, approximately 54 per cent. Individuals ranged from 25 to 119. In this case the rank order correlation between the number of items noted and post-test scores was very small ( $-0.05$ ) and there was no indication that the more thorough pupils did best or that selectivity bore any relation to success. All pupils had noted the key points in the programme, but, clearly, there was no relationship between noting and understanding them. Two conclusions might be drawn: First, that small-step programmes work well for 12-13-year-olds or, second, that the result might have been quite different if the text group had received some form of systematic training in note-taking prior to the experiment. Further work is necessary to decide between these alternatives and to investigate the question of how such training might best be given, for it is a neglected aspect of school work.

The use of the efficiency index invites discussion. The rationale for such a measure is that it gives an indication of success per unit of time taken on the programme. Goldbeck and Campbell (1962) have pointed out that the use of the index is suspect (a) unless one condition of learning is superior to the other in both the amount learned and learning time, and (b) where the relationship between learning time and performance is non-linear. In neither of our two experiments was condition (a) fulfilled. Linearity of regression was tested on the data in Experiment 2 (there were insufficient subjects in Experiment 1) and the correlation ratio was calculated. This achieved a significant value of 0.293 which was greater than the zero order correlation coefficient of  $-0.01$ , indicating a non-linear relationship between time and test score. A linear relationship might be expected in the case of short programmes, but in long ones, each additional increment of time is affected by variations in fatigue, motivation and the like. For this reason, it is suggested that an analysis based on the post test scores is likely to be more reliable. One can conclude that this group of 12-13-year-old pupils derived greater benefit from working through programmes rather than texts.

It would be a crucial factor in favour of programmed learning if it could be shown that material learned in this way could be retained better over a long period of time than that learned from text books. Such a hypothesis could not be supported by the data from these experiments as there were no significant differences between groups in the retest.

In a recent investigation by Alter (1962), retention was tested after 2, 4, 6, 8, 10, 12 and 30 weeks and found to decrease uniformly. This may account for the difference found in the two experiments here. In the second experiment where neither group showed a significant loss, retention was measured after six weeks. In the first, after ten weeks, the scores of both groups had declined significantly.

As previously explained, the tests used in both investigations were designed to test subjects' ability to recall, recognise and transfer their knowledge. The fact that the first investigation showed differences between problems and other items whilst the second showed no differences, suggests that the algebra programme was more effective than the statistics programme in enabling subjects to transfer what they had learned. On the other hand, it was easier to construct a 'problems' section for the statistics test than for the algebra test as the statistics programme covered a wider range of topics. This kind of evidence would be useful in revising a programme, for, if transfer is desired, then the conditions for generalisation to be achieved must be built into the programme.



The fact that the programme group performed better on all types of item than the text group in the second experiment simply confirms the general superiority of this group and suggests the conclusion that under the conditions of this experiment, programmed learning is more beneficial than learning from texts for younger pupils. It has been argued that if the younger pupils had been trained in techniques of note-taking prior to the experiment, this result would not have been obtained. This is an interesting and important question which raises the whole question of training in study skills at grammar school. It is hoped to investigate this topic further.

Restrictions must be placed on the extent to which the findings from the two investigations can be generalised. First, the number of subjects involved was small and replication would be necessary to establish the results conclusively. Secondly, the texts used cannot strictly be compared with ordinary textbooks as they retained some of the features of the original programme and the information was presented in the same sequences. Nevertheless, the results throw light on some important issues. They indicate that the advantages to be gained from using programmes which require a continuous, guided response are likely to vary according to the maturity of the students and the level of skill which they bring to learning from the printed word. In circumstances where this skill is already highly developed, programmed texts which do not require a continuous response will do just as well. As far as younger and less skilful learners are concerned, this study suggests that they are likely to derive some positive benefit from responding at every step. They will also retain their knowledge satisfactorily over a fairly long period of time.

ACKNOWLEDGMENTS.—The writers wish to express their thanks to Dr. J. Annett, for permission to make use of his statistics programme and to the Headmaster and mathematics staff of Ecclesfield Grammar School for their co-operation in providing facilities for the second experiment.

#### V.—REFERENCES.

- ALTER, M. (1962). *Retention in Programmed Instruction*. Tech. Report 620917, 1962. New York: Center for Programmed Instruction.
- EVANS, J. L., GLASER, R., and HOMME, L. E. (1960). A preliminary investigation of variation in the properties of verbal learning sequences of the 'teaching machine' type. See Lumsdaine, A. A., and Glaser, R., 1960.
- FRY, E. B. (1960). A study of teaching machine response modes. See Lumsdaine, A. A., and Glaser, R., 1960.
- GOLDBECK, R. A. (1960). *The Effect of Response Mode and Learning Material Difficulty on Automated Instruction*. American Institute for Research, Tech. Rep. No. 1.
- GOLDBECK, R. A., and CAMPBELL, V. N. (1962). The effects of response mode and response difficulty on programmed learning. *J. Educ. Psychol.*, 53, pp. 110-119.
- HOLLAND, J. G. (1960). Teaching machines: An application of principles from the laboratory. See Lumsdaine, A. A., and Glaser, R., 1960.
- LUMSDAINE, A. A., and GLASER, R. (Ed.) (1960). *Teaching Machines and Programmed Learning*. Washington, D.C., Department of Audio-Visual Instruction, N.E.A.
- ROE, A., et al. (1960). *Automated Teaching Methods Using Linear Programs*. Dept. of Engineering, University of California.
- SIEGEL, S. (1956). *Non-parametric Statistics for the Behavioural Sciences*. New York: McGraw-Hill.
- SILVERMAN, R. E., and ALTER, M. (1961). *Response Mode, Pacing and Motivational Effects in Teaching Machines*. New York: U.S. Naval Training Device Center.
- SKINNER, B. F. (1958). Teaching machines. *Science*, 128, pp. 969-977.

(Manuscript received 20th April, 1963.)



# THE EFFECT OF COUNSELLING ON THE ACHIEVEMENT OF HIGH-ABILITY PUPILS

By

GEORGE SHOUKSMITH

(University of Canterbury, New Zealand)

AND

J. W. TAYLOR

(Teachers' Training College, Tonga)

**SUMMARY.** The under-achievement of high ability children is a growing and important problem. This paper reports the effect of counselling on the level of educational attainment of high ability under-achieving intermediate school pupils. Three matched groups of under-achieving pupils were chosen, known as the Counselling, Placebo and Control Groups, respectively. The Counselling group received non-directive counselling following the administration of a battery of tests. The Placebo group received the same tests but no counselling. The Control group received neither tests nor counselling. The results showed a significant difference in the achievement scores of the Counselling group over both control groups in four of the six tests. In the other two tests, differences approaching significant levels were obtained. This leads to the conclusion that counselling does have a positive effect on under-achieving pupils.

## I.—INTRODUCTION.

It has been common for educational administrators to concentrate special teaching resources on children of low mental capacity. The often unstated belief is that children of superior and high ability are quite capable of developing adequately and fully without any special attention being given to them. Gradually, however, the realization has come that these individuals who should be making significant contributions to society, are, in fact, not doing so. The situation is, as Havighurst (Havighurst, *et al.*, 1955) suggests, that "at least half of our best human material is not developed to anywhere near capacity." This lack of development to capacity is termed 'under-achievement' and the present paper reports an experimental approach to the problem of overcoming under-achievement in a group of children of above-average ability.

## II.—PREVIOUS RESEARCH.

Much of what has been written on under-achieving refers to causes, and little to possible cures. Causes arising from community pressures for example, have been discussed by Mead (1954) whilst Hollingworth (1942) sees mismanagement and inept handling by parents as a major cause of under-achievement. Such incorrect parental attitudes as well as community and social pressures are seen by Strang (1951) to produce a form of maladjustment which, in turn, produces the under-achievement. This conception of the under-achieving child as being emotionally disturbed is supported by a number of investigations (e.g., Barrett 1956, Burns 1949) and is an important factor to be considered in any attempt to overcome the problem. For various reasons, gifted children find difficulty in adjusting and as Strang (1951, *op. cit.*) points out this leads to misdirected ability. Emotional blocks, she suggests, prevent gifted children from realizing their potential, and these blocks are often accompanied by feelings of inferiority or inadequacy and by unsatisfying human relationships. Thus, she regards a school atmosphere of friendly acceptance and encouragement as necessary to produce full pupil participation.



The influence of the school and particularly of the teacher in promoting good working habits has been stressed by many authors (Miles 1956, Louttit 1936). How the teacher is to promote better work habits and better performance with the under-achiever, has received much less attention. Farwell and Peters (1956) are amongst the few authors who specifically state the need for guidance and counselling with under-achievers. Causes originating with the teachers, who may place unnecessary limits on the child, in the school and in the home produce apathy they assert, and this can best be overcome by counselling. Although not speaking specifically of counselling, Cutts and Moseley (1957) referring to the pupil-teacher relationship, point out that "Occasionally . . . a child becomes confidential and tells a great deal about himself, his troubles and worries, his interests and hobbies, his aims and ideas . . ." Such a relationship is the basis of counselling and as the authors conclude, this relationship leads to the pupil wanting to work for you.

### III.—THE COUNSELLING HYPOTHESIS.

Since under-achievement appears often to be the result of personal mal-adjustment of some sort, it would follow that counselling of under-achievers should offer a possible solution to the problem. All teachers would agree that friendly contact with pupils, and individual attention to their wants is of value and can effect a change of attitude to work and study. Furthermore, it is suggested (Shouksmith 1960) that personal interest and attention shown by someone not emotionally involved is the basis of good counselling.

With these considerations in mind, it was decided to test the hypothesis that a short period of counselling would help children of high intelligence, but low achievement, to overcome some of the personal problems which affect their school work. The counselling technique used, was made as simple as possible so that if the method proved successful it could readily be used by general teachers without a long period of training in counselling. Counselling was generally non-directive and the counsellor acted as a sounding board, a sympathetic listener who showed a personal interest, who encouraged the subject to seek his own solutions and who offered suggestions only when an impasse seemed to have been reached and help was sought.

The group of under-achievers chosen for the experiment came from Form I of a N.Z. intermediate school. The N.Z. intermediate school caters for 12 and 13-year-old children in the main and was chosen as the population to be sampled because previous experience had shown that the more intelligent of this age group were capable of understanding the mechanics of personal problems and of working out feasible solutions to many of their own problems. Furthermore, practical experience in teaching these pupils showed that they often needed support and help in tackling problems which appeared too large to tackle alone.

### IV.—METHOD.

Three groups of twelve were chosen from 331 first year children in a N.Z. intermediate school. These groups were formed by assigning members of trios matched as closely as possible on the basis of I.Q., age, sex and achievement scores, one to each group. All thirty-six pupils were chosen from a high ability, under-achieving sample of the intermediate school population. An OTIS Intermediate Test was used to screen out those pupils of above 116 I.Q. From this group those pupils who were under-achieving on a criterion of falling two deciles below expectancy, based on the I.Q. figure, were further screened out. The tests used to measure achievement were the A.C.E.R. Reading-Arithmetic Battery.



These three groups were then used as follows :

*Group 1—Experimental Counselling Group.*

This was the group to whom close individual attention and counselling was given. In addition, certain additional tests were administered, mainly in order to gain information upon which subsequent counselling sessions could be based. Each child was interviewed individually on some occasions and in small groups on others, for periods ranging from twenty to forty minutes, at approximately fortnightly intervals, for a period of six months. The number of interviews varied with the seen needs of the children concerned. The children were made aware of the fact that further interviews were available at any time upon request and some asked for additional sessions with the Counsellor.

*Group 2—The 'Placebo' Group.*

It might be argued that the administration of additional tests, indicating that special interest was being shown in a group, might have an effect on the children concerned, irrespective of counselling received. To control this aspect of the experiment, a second group was formed who received all the additional tests given to the Counselling Group but no counselling.

*Group 3—The Control Group.*

This group's identity was known only to the experimenters. All Form I pupils were given the initial screening and final appraisal tests and this group was formed on paper only when their results on the various measures were extracted for comparison purposes.

The design of the experiment may be seen in diagrammatic form in Table 1.

TABLE 1  
TIME-TABLE OF EXPERIMENTAL PROCEDURE FOR THREE GROUPS.

	Group 1 Counselling Group	Group 2 'Placebo' Group	Group 3 Control Group	All Other Form I Pupils
INITIAL SCREENING :				
Pre-Experimental Period	OTIS A, B	OTIS A, B	OTIS A, B	OTIS A, B
Pre-Experimental Period	ACER Achievement Battery	ACER Achievement Battery	ACER Achievement Battery	ACER Achievement Battery
Months 1 and 2 . . . . .	Sociogram	Sociogram	Sociogram	Sociogram
	Parents Interviewed			
	W.I.S.C.	W.I.S.C.		
	California Personality Test	California Personality Test		
	Sentence Completion	Sentence Completion		
	Health Survey	Health Survey		
Months 2-6 . . . . .	Counselling		Sociogram	Sociogram
Month 6 . . . . .	Sociogram	Sociogram		
	Parents Interviewed			
FINAL APPRAISAL :				
Beginning Month 7 . . . . .	ACER Battery	ACER Battery	ACER Battery	ACER Battery

Counselling was carried out in the school staffroom and sessions whether individual or group were kept as informal as possible. In individual counselling the session was usually initiated by the posing of a question for the subject to take up, after which beginning a generally non-directive approach was taken. Occasionally, when an impasse was reached, the counsellor took a more directive line, however. Encouragement and help were also given whenever it appeared that the child was in need of direct support. During the course of the experiment, several group discussions were held. Children with common problems were chosen from the experimental group for such sessions. For example, those children who had difficulty in getting on with siblings constituted one such group. Each counselling session, whether individual or group was concluded by the counsellor giving a generalized summary of the discussion and making arrangements for future meetings.

The parents of children in the experimental group were seen twice, once at the outset of the experiment and again at the end of the programme. At the initial meeting, the programme was explained. As much background information as possible was collected and parents questions answered. At the final meeting, the problems of their own children were discussed with the various parents. If a parent asked for help or guidance, this was given whenever possible.

#### V.—RESULTS.

The Means and Standard Deviations for initial and final tests are summarized in Table 2. The data were analysed by Analysis of Variance to test the significance of difference among means at initial testing and again at final testing.

TABLE 2

SUMMARY OF RESULTS FOR THE THREE GROUPS ON INITIAL AND FINAL TESTING.

Test		Group 1 Counselling Group		Group 2 'Placebo' Group		Group 3 Control Group	
		Initial	Final	Initial	Final	Initial	Final
Addition .....	Mean	3.16	5.45	3.67	4.18	3.5	3.77
	s.d.	2.0	1.41	0.85	1.37	1.35	1.35
Subtraction .....	Mean	4.0	6.1	4.16	4.72	4.25	4.27
	s.d.	1.5	1.67	0.96	1.6	1.27	1.05
Multiplication .....	Mean	4.0	6.82	4.0	4.72	4.25	4.1
	s.d.	1.3	1.44	2.22	1.28	1.25	1.6
Division .....	Mean	4.25	6.18	4.5	5.27	4.75	5.0
	s.d.	1.6	1.45	1.7	0.86	1.3	1.8
Word Knowledge ..	Mean	5.5	7.09	5.08	5.72	5.67	6.0
	s.d.	0.64	0.69	1.06	0.75	1.06	1.28
Reading for Mean- ing .....	Mean	5.84	6.27	5.25	6.09	5.16	5.54
	s.d.	0.64	0.75	1.06	0.58	1.06	0.99



The results of this analysis are shown in Table 3 which shows the F-ratio for variance among means on the initial test as  $F_x$  and the ratio for variance among means on the final test as  $F_y$ . In each case, with d.fs. respectively 2 and 30 an F ratio of 3.32 is significant at the .05 level and one of 5.39 is significant at the .01 level.

TABLE 3

RESULTS FROM ANALYSIS OF VARIANCE TAKEN AMONG MEANS ON INITIAL TEST ( $F_x$ ) AND ON THE FINAL TEST ( $F_y$ ) FOR THE ACHIEVEMENT BATTERY.

	$F_x$	$F_y$
(1) Addition .....	.11	3.55
(2) Subtraction .....	.012	4.19
(3) Multiplication .....	.352	9.508
(4) Division .....	.11	1.44
(5) Reading for Meaning ..	.38	2.2
(6) Word Knowledge ....	1.59	3.82*

\* Corrected by analysing co-variance, because of slight relationship shown in initial means.

It can be seen from Table 3 that in no case was there a significant F-ratio among initial means, so that we may conclude that the groups were equally matched for achievement at the outset. In the final test, Division and Reading for Meaning show no significant variation but for all other measures a significant F-ratio was produced and individual differences between measures were tested for significance. Table 4 gives the result of this series of t-tests and shows the achievement measures for which significant differences between groups were found.

TABLE 4

SUMMARY OF T-TEST RESULTS FOR DIFFERENCES BETWEEN GROUPS ON FINAL TESTS.

(Figures in body of table show significance level.)

Pairs of Means	Achievement Measures*			
	Add.	Sub.	Mult.	W.K.
Test and Placebo .....	.05	.05	.01	.01
Test and Control .....	.01	.01	.01	.01
Placebo and Control .....	N.S.	N.S.	N.S.	N.S.

\* Division and Reading for Meaning were not included in the analysis since F-ratios for these tests were not significant.

Results for individual subjects revealed that eight of the Counselling Group subjects were no longer classed as under-achievers, although in some tests these subjects were not completely satisfactory. Three pupils, however,

reached an entirely satisfactory standard. On the other hand, all of the Placebo Group and ten out of the twelve Control Group were still classed as under-achievers.

Equally as important are various other changes noted in the Counselling Group. The sociograms, for example, showed improved peer acceptance for the Counselling Group; these children were either more popular or less unpopular than they had been formerly. Class teachers found differences emerging too. They reported that children from the Counselling Group were co-operating more readily and showed improved social adjustment. The Placebo and Control Group showed no such similar improvements when their results were compared at the completion of the investigation. Finally, both teachers and parents reported a change of attitude towards school work in the Counselling Group.

#### VI.—DISCUSSION.

The results shown in Tables 2, 3 and 4 indicate quite clearly that general counselling of a non-directive sort has a significant effect on the level of educational attainment of a group of high ability but under-achieving intermediate school pupils. In four of the six achievement tests used, final means for the Counselling Group differed significantly from the other two groups and in the remaining two, although analysis of variance did not permit a test of individual means, the means of the Counselling Group are, nonetheless, higher as can be seen from Table 2. The individual results, which show a large proportion of the Control Group children moving out of the under-achieving category also support this conclusion.

It is true that the improvement was not sufficient to bring all subjects in the Counselling Group up to a normal level, but this could hardly be expected. The duration of the experiment and hence the extent of the counselling given, was limited by practical necessity. A longer period of counselling might well have produced even greater improvements. Also, in many of the children, the under-achievement was of long standing and involved the loss of a good deal of basic material. Sometimes, the reason for a child under-achieving seemed to be not the more generally found personal mal-adjustment and, therefore, counselling could hardly be expected to succeed. Child I in the Counselling Group, for example, showed practically no improvement, particularly in the arithmetic tests. This child, however, appeared to have a completely normal social and personal adjustment but had missed over a year of schooling in short periods since she started school.

In any study of this nature it is impossible to control all variables and, in particular, one other factor outside the experimenters' control, influenced the results. This was the attitude of the parents of the Counselling Group children. It was noteworthy that the seven parents who co-operated most willingly in both the first and second meetings and who expressed most concern and were eager to help, had children who profited most from the counselling. On the other hand, there were three parents who even after long discussions at the outset of the experiment, were unwilling to let their children participate, and who failed to attend the second meeting. Significantly, their children seemed to profit least.

The Placebo Group was included initially to test the possibility that the mere fact that someone was paying more attention to these children and "appeared to be doing something about them," would be sufficient to cause an improvement, without counselling being necessary. It seems reasonable to



suggest that the administration of the tests to these children made them aware that they were under special consideration of some sort and this, against the background of the general awareness of the school that a programme of testing and counselling was being undertaken, does seem to have had a beneficial effect. Table 2 shows that on the final assessment the mean scores of the Placebo group on all achievement measures were higher than those of the Control Group. Reference to Table 4, however, shows that in no case was the Placebo Group's improvement a statistically significant one. Furthermore, on all tests the means of the Counselling Group are still higher. Special attention without counselling is not sufficient to overcome under-achievement.

From the practical standpoint, the results of this experiment suggest that the present school programme should be extended to include counselling with under-achieving pupils, particularly those of high ability. The nature of the counselling used in the present experiment was made as simple as possible, in the hope that such procedures could be applied by teachers with a minimum of psychological training. A short training course in the specialized procedures involved could be given to suitable classroom teachers, enabling counselling programmes to be introduced on a wide scale. The significant improvement of children in the Counselling Group in this experiment, indicates that time spent this way should be profitable and rewarding both to child and teacher.

#### VIII.—REFERENCES.

- BARRETT, H. G. (1956). *Under-achievement, a Pressing Problem*. Ball, Ontario : Second ary School Teachers' Federation.
- BURNS, C. L. C. (1949). Maladjusted children of high intelligence. *Brit. J. Educ. Psychol.*, 19, pp. 137-141.
- CUTTS, N. E., and MOSELEY, N. (1957). *Teaching the Bright and the Gifted*. New York : Prentice Hall Inc.
- FARWELL, G. F., and PETERS, H. J. (1959). *Guidance Readings for Counsellors*. Chicago : Rand, McNally and Co.
- HAVIGHURST, R. J., STEVENS, E., and DE HAAN, R. F. (1955). A survey of the education of gifted children. *Supplementary Educational Monograph*, No. 83. Chicago : University of Chicago Press.
- HOLLINGWORTH, L. H. (1942). *Children above 180 I.Q., Part III*. New York : World Book Co.
- LOUTTIT, C. M. (1936). *Clinical Psychology*. New York : Harper Bros.
- MEAD, MARGARET (1954). The gifted child in American culture today. *J. Tech. Ed.*, 5, pp. 211-214.
- MILES, CATHERINE in CARMICHAEL, L. (Ed.) (1956). *The Gifted Child in Manual of Child Psychology*. New York : John Wiley and Sons.
- SHOUKSMITH, G. (1960). Developing clinical psychology in New Zealand. *Hospital J. of A.P.A.*, 11, pp. 41-43.
- STRANG, RUTH, in WITTY, P. (Ed.) (1951). *Mental Health of Gifted Children, in The Gifted Child*. Boston : Health and Co.

(Manuscript received 22nd January, 1963.)

# A STUDY OF SOME COGNITIVE AND OTHER DISABILITIES IN BACKWARD READERS OF AVERAGE INTELLIGENCE AS ASSESSED BY A NON-VERBAL TEST

By K. LOVELL, D. SHAPTON AND N. S. WARREN

(*University of Leeds*)

**SUMMARY.** The application of a Sentence Reading Test and a Non-Verbal Intelligence Test to the whole of the third year pupils in twenty-two junior schools gave results for 1,205 children. Of the 139 who had a reading quotient of  $<80$ , seventy had standardized non-verbal scores  $\geq 90$ . Fifty of these seventy backward readers were paired, individually, with average to good readers, matched for non-verbal test score, social class, sex and school. These two groups undertook a number of individual tests. The backward readers showed no deficit in performance on an oral test of language structure, but a poorer performance on a variety of tests involving spatial relationships and left-right discrimination; also a greater 'rotation effect' on a test involving the copying of abstract designs.

## I.—INTRODUCTION AND PROBLEM.

THERE is an increasing literature dealing with the aetiology and concomitants of backwardness in reading. It is not our purpose to review this literature here; recent examples of relevant papers are those of Vernon (1962), Crookes and Greene (1963), Ingram (1960), while the very recent formation in Great Britain of 'The Association for Word Blind Children' indicates the growing interest in this general field.

The present study arises out of an earlier one. In an investigation commenced in the Autumn of 1962, the N.F.E.R. Sentence Reading Test and the N.F.E.R. Non-Verbal Intelligence Test No. 5 were given to the whole of the third year pupils in twenty-two junior schools. The results of children who had had a change of junior school were ignored, and this left data on 1,205 pupils (643 boys, 562 girls) who had been in the same school since 7+ years of age. Of the 139 pupils (101 boys, 38 girls) who had a reading quotient  $<80$ , sixty-one boys and nine girls had a standardized non-verbal intelligence test score  $\geq 90$ . Moreover, thirty-two boys, but only three girls, had a non-verbal score  $\geq 100$ . These figures suggested the hypothesis that about a half of backward readers (R.Q.  $<80$ ) have average or higher non-verbal intelligence test scores, with a far higher proportion of backward boys than girls possessing average or better non-verbal reasoning skills.

As a check on this hypothesis, data were obtained on a further 1,464 children (768 boys, 696 girls) in the second half of their first year in the junior school. These pupils constituted the whole of the year group of sixteen junior schools in a different area. The tests given were the N.F.E.R. Sentence Reading Test and either the N.F.E.R. Picture Intelligence Test or the Moray House Picture Test No. 1. In this population there were 172 boys and 81 girls with a reading quotient  $<80$ . Of these, ninety-five boys and thirty-one girls had non-verbal intelligence test scores  $\geq 90$ , and thirty-seven boys and ten girls with a non-verbal test score  $\geq 100$ . It will be seen at once that the data obtained from this second group confirmed our hypothesis.



It would appear then that in the ordinary 'run of the mill' backward pupils—and not necessarily those designated by special labels—there may be some specific disability affecting the male more than the female. This may well account for the far greater incidence of reading backwardness among boys than girls that seems to be found in all countries. In the light of this information, we selected from our first group of 1,205 third year junior school children, fifty out of the seventy pupils who had reading quotients  $< 80$ , but who had non-verbal intelligence test scores  $> 90$ . Furthermore, for 96 per cent. of the 1,205 children, details of parental occupation were made available through the co-operation of the head teachers concerned. From the occupations it was possible to get the social class of the parents through reference to the Registrar-General's *Classification of Occupations*, 1960, H.M.S.O., 1961.

Each of the fifty backward readers (forty-one boys, nine girls) was then paired with an average to good reader, the members of each pair being matched for non-verbal intelligence test score, social class, age, sex, while both members of each pair were pupils of the same school. So, one-hundred children were tested individually in a number of tests drawn from the current literature.

## II.—PROCEDURE.

The relevant literature provided an embarrassing selection of tests from which to choose, and the list of tests given below was finally selected as being the most representative and as being most likely to differentiate between the two groups of children. Tests Nos. 1-4 were administered by D.S. and Nos. 5-9 by N.S.W. and K.L.

(1) Watts' 'An English Language Scale' (Watts, 1944). This test measures the progress made by children in mastering, orally, the basic varieties of the English sentence.

(2) A sentence copying test. Ten sentences had to be copied by the child. The sentences contained words of one letter, two letters, etc., up to ten letters, and there were five examples of each word of a given length. The number of errors made was recorded under the headings: reversal of letter, distortion of letter, incorrect letter, wrong or indecipherable word, letter or word omitted, capital letter instead of small letter or *vice versa*, word misplaced.

(3) Benton (1959) tests of left-right discrimination. Herman (1959) suggested that directional confusion resulted from the disturbance of the body schema related to impaired functioning of the parietal lobes, although directional confusion is not invariably found in reading disability. Benton (*op. cit.*) has investigated the problem of directional confusion in the field of left-right discrimination and has provided a number of tests (Tables 1 and 2, p. 14 and p. 15). The literature suggests that children with reading disabilities may well show retarded development of left-right discrimination in early years, but that performance is within the normal range by about nine. But there is uncertainty about this, and the tests proposed by Benton were used with our population of children, in spite of the fact that their ages ranged from 9.4-10.4 years. There were twenty-eight tests relating to the child's own body, and sixteen relating to his body and/or the picture of a man (e.g., 'put your right hand on the man's left ear').

(4) Tests involving Axes of References devised by Piaget and Inhelder (1956). It has also been stated by Inhelder (1961) that children with reading disabilities made a poorer performance on these tests than average readers do. The three tests used involved:

- (a) The level of water when the container was tilted at various angles (round and straight sided flasks used).
- (b) The relation of a 'ship' and its 'mast' to the sides of a flask when the latter was tilted at various angles.
- (c) The uprightness of sticks placed on a 'mountain' made of paper and plasticine. A further task was the drawing of houses on the sides of a hill.

(5) WISC Vocabulary test.

(6) WISC Block Designs test.

(7) WISC Object Assembly test.

(8) WISC Coding test (equivalent to Digit-Symbol in the adult form of the test).

(9) Shapiro's test of Rotation (Shapiro, 1962). The 'rotation effect' refers to a tendency to reproduce abstract designs correctly, but, in a mis-orientated position. This test has been shown to yield significant differences between certain kinds of brain injured patients and normal adults. Six designs were used and the test was carried out by the subjects drawing each design on the fresh page of a pad, using a blue pencil.

No tests of articulation were given because of our findings in another study, the results of which it is hoped to publish later. In that investigation the whole population of a small primary school (100 children) underwent a very thorough series of tests of articulation given by a member of the Department of Phonetics, Leeds University, who had previously been a school teacher. It was found that up to 8+ years of age, the number of errors of articulation made by children whose reading quotients were  $< 90$ , was significantly greater than the number of errors made by pupils whose reading quotients were  $\geq 90$ . But, in the third and fourth years of the junior part of the school, no significant difference in the number of errors of articulation could be found between the two groups of readers.

### III.—THE RESULTS.

The reading quotients of sixteen of the fifty backward children were  $< 70$ , but in order to calculate a rough mean for the backward group, these quotients were counted as 70. The distribution of non-verbal intelligence test scores was asymmetrical with most of the scores bunched between 95 and 110. Accordingly, the range of non-verbal scores is given but not their standard deviation.

The raw rotation scores were subjected to a logarithmic transformation, because of their negative skewness, in an attempt to normalize them for the purpose of testing for a significant difference in mean scores between the two groups.



TABLE 1

INITIAL DATA RELATING TO THE MATCHED GROUPS OF CHILDREN.

	Backward Group	Non-Backward Group
Mean C.A. ....	9.8 yr.	9.8 yr.
Mean R.Q. ....	74	106
Range of R.Qs. ....	<70-79	$\sigma=6.3$
Mean Non-Verbal Test Score ....	103	103
Range of Non-Verbal Test Scores ....	90-131	93-127
Sex ....	41 boys 9 girls	41 boys 9 girls
Social Class ....	II 1 III 22 IV 21 V 6	II 1 III 22 IV 21 V 6

TABLE 2

DATA OBTAINED FROM TESTS 1-4 FOR MATCHED GROUPS.

Test No.	Test	Backward Group	Non-Backward Group	Difference Significant at 1 per cent. level
1	Score on Watts' Test in Years....	M=8.91 $\sigma=.35$	M=9.07 $\sigma=.28$	No
2	Total number of Errors in Copying Test .....	448	238	Yes
	Number of Errors involving Distortions, Incorrect Letter, Wrong or Indecipherable Word .....	271 177	88 150	Yes No
3	Number of Correct Responses on Benton Tests : Own Body .....	1215 510	1273 527	Yes No
4	Axes of References (Results for Three Tests Combined) .....	20 93 37	5 81 63	Yes
	Number at Piaget Stage IIA ..			
	Between IIA and IIIA, inclusive .....			
	Better than IIIA .....			

TABLE 3  
DATA OBTAINED FROM TESTS 5-9 FOR MATCHED GROUPS.

Test No.	Test	Backward Group	Non-Backward Group	Difference Significant at 1 per cent. level
1	Scaled Score on WISC Vocabulary	M=7.08 $\sigma$ =1.90	M=9.42 $\sigma$ =1.84	Yes
2	Scaled Score on WISC Block Designs .....	M=10.28 $\sigma$ =2.80	M=11.52 $\sigma$ =2.44	No, but significant at the 2 per cent level
3	Number of completed Designs Incorrect because of Reversals Number of Completed Designs without Reversals .....	47 135	28 193	Yes
4	Scaled Score on WISC Object Assembly .....	M=8.12 $\sigma$ =2.48	M=8.92 $\sigma$ =2.19	No
5	Scaled Score on WISC Coding ..	M=9.56 $\sigma$ =2.12	M=10.94 $\sigma$ =2.68	Yes
6	Raw Score on Shapiro Rotation Test (Mean Rotation on Six Designs for each Subject) .....	M=14.2°	M=10.2°	Yes, after Logarithmic Transformation of Raw Scores.

#### IV.—DISCUSSION.

While the tables show that there are significant differences between the two groups of children on certain tests, it became abundantly clear when making a detailed examination of the performance of individual children that there was a good deal of overlap between the performance of the members of some pairs. Thus, sometimes the backward reader made a better score on a test than his partner. Indeed, the backward child sometimes gave a poor response on one test, sometimes on another, but not consistently on a number of tests. However, the upshot of our findings may be summarized as follows:

(a) On the Watts test of language, there is very little difference between the mean scores of the two groups. In other words, the poor readers are able to handle the basic varieties of the English sentence almost as well as normal readers providing the task is an oral one. Note that the mean language age for each group is somewhat below the mean chronological age. This would be expected from the distribution of social class within the groups.

(b) The backward readers, on the average, made about twice as many mistakes as their partners did on the test of sentence copying. But very many backward readers made few errors, and some made none at all. The errors found particularly frequently among the poor readers were distortions, incorrect letter, wrong or indecipherable word. The other errors were found with about the same frequency in both groups.



There tended to be a sharp increase in the number of errors made by the backward readers when the word was of four letters or more, whereas the corresponding increase of frequency of error occurred in the normal readers with a word of seven or more letters. Furthermore, judged against Burt's quality scale of handwriting (Burt, 1921), the backward readers obtained far lower ratings than their partners when the written sentences were rated independently by D.S. and another person, the latter person not knowing if the script belonged to a poor reader or not. Twenty-seven of the backward readers got below average ratings against eight in the other group.

(c) On the Benton tests the poor readers did less well than their partners when left-right discrimination of their own bodies was involved. But it must be stressed that many of the poor readers make no more errors than their partners and some make none at all. For both groups considered together, there was roughly a 90 per cent. correct response. All we can say is that these tests were badly performed by some pupils with reading disability.

But the tests that involved left-right discrimination in relation to the picture of another person, failed to differentiate between the two groups. These tasks were harder for all the pupils and the success ratio dropped to 60/65 per cent.

(d) The tests involving Axes of Reference showed clearly that the responses of the backward readers were far more frequently at the lowest stages, and far less frequently at the highest stages, compared with average readers. This suggests that reading disability is frequently linked with difficulty in building up a frame of reference involving spatial relationships. Inhelder's view was thus confirmed.

(e) On the WISC Vocabulary test the poor readers did much less well than the other group. This was expected. Bearing in mind the other evidence, it is suggested that the score is lower, on the average, because the child has a reading disability, and not the other way round.

(f) There was a small but significant difference in mean score in favour of the better readers on the WISC Block Designs test. This confirms what other workers have found. More important though, was the number of reversals—i.e., a rotation through  $90^\circ$  or  $180^\circ$  of the blocks divided diagonally into red and white—in those designs that were finally completed as a square of four or nine cubes. The backward readers completed 182 designs within the time limits of which forty-seven were incorrect because of reversals. The corresponding numbers for the other group were 221 and 28. These differences are highly significant.

(g) There was a small but significant difference in the means of the WISC Coding Test in favour of average readers, thus confirming the findings of De Hirsch (1954), and Drew (1956). But there was no significant difference in the means of the groups on the WISC Object Assembly test.

(h) The backward readers scored significantly higher rotation scores, compared with the other group, on the Shapiro test. Shapiro, *et al.*, suggests that the increased rotation results from restriction or disturbance of visual input caused by brain lesions leading to visual field defect or 'ocular-motor' defects. Our poor readers performed relative to his group of normals. If mental brain injured group performed relative to his group of normals. If Shapiro, *et al.*, are correct, one may suggest that reading disability in many children of average non-verbal intelligence test score results from some brain dysfunction which, in turn, gives limited ability in the decoding and encoding processes involved in reading, and in some cases, restriction or disturbance of visual input.



Summing up, we may say that in a group of fifty 'ordinary' backward readers, selected from a group of 1,205 third year junior school pupils, and matched with average and better readers for non-verbal intelligence test score, social class, age, sex, and school, there was :

- (i) No deficit in performance on an oral test of language structure but a poorer performance on a vocabulary test.
- (ii) A poorer performance on a variety of tests involving spatial relationships and left-right discrimination ('directional confusion').
- (iii) A large number of errors in the copying of words: particularly certain types of error.
- (iv) A greater 'rotation effect' on a test involving the copying of abstract designs.

In addition, there is a very marked sex discrepancy in the number of boys and girls who are backward readers, but of average or better thinking skills. In the light of this evidence, it seems very unlikely that methods of teaching, sociological factors and emotional problems can account for all reading backwardness. It is reasonable to suggest that disability in reading may be but one symptom of brain dysfunction. It could be genetic in origin. This would also explain why Lovell (1962, 1963) has found that the majority of children who are still backward at 9 years of age, make only limited progress in reading under remedial teaching almost regardless of I.Q. In the case of those who do make rapid and sustained progress, it could be that signals are re-routed at some stage or new areas take over from areas not functioning properly, or there is some maturational effect at work. Further research in this general field is urgently required.

There is also an important educational implication for the classroom. Many backward readers, particularly boys, have considerable powers of thought, are capable of a good level of discussion at the oral level, and have the ability to assimilate experience and make progress in many areas of knowledge providing reading and writing make limited demands upon them.

#### V.—REFERENCES.

- BENTON, A. L. (1959). *Right-Left Discrimination and Finger Localization*. New York: Hoeber-Harper.
- BURT, C. (1921). *Mental and Scholastic Tests*. London: Staples.
- CROOKES, T. G., and GREENE, M. C. L. (1963). Some characteristics of children with two types of speech disorder. *Brit. J. Educ. Psychol.*, 33, 31-40.
- DE HIRSCH, K. (1954). Gestalt psychology as applied to language disturbances. *J. Nerv. Ment. Dis.*, 120, 257.
- DREW, A. L. (1956). A neurological appraisal of familiar congenital word-blindness. *Brain*, 76, 440.
- HERMANN, K. (1959). *Reading Disability*. Copenhagen: Munksgaard.
- INGRAM, T. T. S. (1960). Paediatric aspects of specific developmental dysphasia, dyslexia and dysgraphia. *Cerebral Palsy Bulletin*, 2, 254-277.
- INHELDER, B. (1961). Contribution de la méthode génétique à l'étude de quelques phénomènes pathologiques de la pensée. Paper read at the XIV International Congress of Applied Psychology.
- LOVELL, K., JOHNSON, E., PLATTS, D. (1962). A summary of a study of the reading ages of children who had been given remedial teaching. *Brit. J. Educ. Psychol.*, 32, 66-71.
- LOVELL, K., BYRNE, C., RICHARDSON, B. (1963). A further study of the educational progress of children who had received remedial education. *Brit. J. Educ. Psychol.*, 33, 3-9.
- PIAGET, J., and INFELDER, B. (1956). *The Child's Conception of Space*. London: Routledge and Kegan Paul.
- SHAPIRO, M. B., BRIERLEY, J., SLATER, P., BEECH, H. R. (1962). Experimental studies of perceptual anomaly VII. *J. Ment. Sci.*, 108, 655-668.
- VERNON, M. (1962). Specific dyslexia. *Brit. J. Educ. Psychol.*, 32, 143-150.

(Manuscript received 3rd April, 1963.)



# CLASS SIZE AND CO-ORDINATED INSTRUCTION\*

By JOHN P. DE CECCO  
(*San Francisco State College*)

**SUMMARY.** The hypotheses tested were that there would be significant differences in student achievement and preference for section, attributable to class size and co-ordinated instruction. Students were freshmen psychology students ( $N=682$ ). Students were randomly assigned to three types of instructional groups: large-experimental (LE), small-experimental (SE), and small-control (SC). For the LE and SE groups instructional materials and assignments were co-ordinated. The major criterion measure tested for cognitive change beyond the acquisition of information. In the main, there were no significant differences in achievement or preferences. However, students consistently prefer small classes ( $p < .001$ ) when asked directly about this factor, although other responses are not consistent with this expression of attitude.

## I.—INTRODUCTION.

THERE is considerable justification for the continuation and extension of the research on class size. First, references to the 'vast' research literature on class size are many times more frequent than the actual number of studies. Few studies have been carried on by educational psychologists and other professional educational researchers or published in research journals. Secondly, the criteria measures frequently employed have limited validity for the measurement of cognitive behavior change in general psychology beyond the recall or recognition of relevant course information. Thirdly, almost none of the previous studies used a random assignment of students for the various instructional treatments. Fourthly, co-ordination of instructional materials frequently did not extend beyond the use of a common text and the criterion measure. The scheduling of content of all major assignments, all instructional films, all quizzes and examinations, as well as the textbook and criteria measures, were co-ordinated in this study.

The most extensive studies of class size were conducted in 1920 at the University of Minnesota (Hudelson, 1928). Most differences reported in these studies were not significant. However, the majority of significant differences in student achievement favored larger classes. Small classes ranged from twelve to sixty in size, while large classes ranged from thirty-five to 150. Most criterion measures were tests of knowledge acquisition.

Before World War II, most of the experimentation on class size was conducted at the primary and secondary school levels, partly in response to financial problems caused by the depression (Amundson, 1954). In the *post-bellum* years, these investigations have been extended to colleges and universities as student enrollment swells. These later studies have more or less successfully attempted to relate class size to cognitive change beyond the acquisition of information, as described in Bloom's *Taxonomy of Educational Objectives*. Rohrer (1957) compared large classes of 309 and 332 students with small classes of twenty-three and thirty-one. The criterion measure was the Co-operative

\* Grateful acknowledgement is made to the Psychology faculty of San Francisco State College, who authorized the project, and to the several instructors who participated in the experiment. Special acknowledgement is made to Professors Dan Adler, Bela Baker, and Jerome Podell, who acted as consultants and to Carol Robinson, my research assistant.



American Government Test which requires students to apply the principles and process learned. There were no significant differences in achievement for large and small groups. Cammorosano and Santopolo (1958) compared achievement in large and small classes in economics, political science, and sociology. The large classes numbered only sixty while the small classes (almost as large) numbered about thirty. Besides short-answer and multiple-choice tests to measure the acquisition of subject matter, they used tests of 'social awareness' to assess ability to identify influential public events and to develop 'principled syntheses of social outlook.' Again, no significant differences were reported. Nelson (1959) used classes in economics. Large sections ranged from eighty-five to 141 students and small sections from sixteen to twenty students. An attempt was made to control for instructor variability by using only two instructors in each of three semesters and only one of these instructors in the fourth semester. Students were matched on the basis of sex, year or class in college, and subject major, but not for previous course knowledge, I.Q., and grade-point averages. A criterion measure composed of multiple-choice and true-false items showed no significant differences for the various treatments. Macomber and Siegel (1960) undertook a major study on class size in many disciplines, although their major purpose was the assessment of the effectiveness of closed-circuit televised instruction. For the two courses in basic psychology there were seven large groups ranging from seventy-eight to 205 students and eight small groups ranging from thirty to forty-two students. The following criterion measures were used: objective tests for subject-matter acquisition; Dressel's Test of Critical Thinking in Social Science and his test of Science Reasoning and Understanding; and a test of misconceptions in psychology composed of true-false statements. The Dressel tests have greater validity for the natural rather than the behavioral sciences and true-false items of the misconceptions test reduce its reliability. There were also opinionnaires and scales used to discover student attitude toward the various instructional procedures. Gross and analyses of data indicate no significant differences in all measures of student achievement. However, students in larger sections tended to rate course content somewhat less favorably than did students in small sections. This finding is consistent with those of Lovell and Hanes (1955), where student reaction favored small classes and instructors of small classes but not consistent with an earlier study (Riley, *et al*, 1950) in which no significant differences in student reaction were found. Students were not randomly assigned but rather statistically matched on the basis of aptitude and achievement. This procedure resulted in the elimination of unmatched students who remained members of the various classes.

In the present experiment, we tested the hypothesis that class size would be a significant variable in determining academic achievement for students in small as compared with large classes. We further hypothesized that students (in a subjective evaluation of their experiences) would show a significant preference for one of the class sizes. We also hypothesized that co-ordination of instruction would significantly influence student achievement in the different sections and affect their subjective reaction to the course and to specific course assignments.

*Procedure.* Three types of instructional groups were created: (1) two large experimental groups, one with 127 students, and the other with ninety-seven students (hereafter designated 'LE')\*; six small experimental groups, with a

\* Only those subjects who took both the pretest and post test are included. The estimated rate of drop-out or absence was not significantly different for the various types of groups.



range of thirty-four to eighteen and a mean of twenty-eight students (hereafter designated 'SE'); and ten small control groups, with a range of thirty-five to twenty-two and a mean of twenty-eight (hereafter designated 'SC'). Each section had a different instructor with these exceptions: two SC sections had the same instructor, and two SE sections had the same instructor. Classes were scheduled from 9 a.m. to 3 p.m. to control for possible differences in characteristics of students who register for morning against afternoon classes. The schedule also provided for parallel meeting of experimental and control groups throughout the morning and afternoon.

Instruction in the large (LE) and small (SE) experimental groups was co-ordinated in the following manner. These groups used the same course outline which specified the major course assignments and their sequence. The major assignments included two critical reports and an experimental project. They used the same textbook: Ernest R. Hilgard, *Introduction to Psychology* (New York: Harcourt, Brace and World, Inc., 1963) and the accompanying *Student Guide* which programs most of the major concepts of the textbook. The same films, quizzes and examinations were used. A common grading system was developed for evaluating the major reports and for the assignment of the students' final course grades. Beyond this co-ordination of assignments, sequences, materials, and examinations, no further effort was made to establish uniformity of instructional technique or course content. Instructors in LE and SE sections met weekly to co-ordinate work for the following three class meetings.

The small (SC) control groups were ten of the intact sections of our ongoing General Education course in introductory psychology. Currently, instructors in these groups select their own materials, create their own assignments, and construct their own examinations. Except when knowledge about the experiment may have modified these practices, instructors in the sections had the usual degree of independence in making all course decisions.

There was no special basis for the selection of instructors for the control and experimental groups. However, before assignment was made to an experimental section, it was necessary to secure the consent of the instructor for the co-ordination of his section. The two instructors for the LE sections were selected primarily on the basis of previous experience in teaching large classes. All SE and SC instructors had experience with small classes.

*Subjects.* The subjects in this experiment were first-semester freshmen at San Francisco State College. They registered for a required course in introductory psychology without prior knowledge of the experiment or the type of instructional group to which they were assigned. For each instructional hour throughout the morning and afternoon of which experimental and control groups were scheduled and on the first day of classes, students were randomly assigned to the various groups described above. They were handed cards which designated the room they were to report to. These cards had been randomized by using a Table of Random Numbers.

*Criterion measure.* The chief measure of achievement used was Form A of the *Introductory Psychology Criteria Test* developed at the University of Michigan by the Committee on Criteria for the First Course which has been tied in with a project on Evaluating the Effectiveness of Instruction in Higher Education, supported by the Co-operative Research Branch of the U.S. Office of Education. Eight objectives for the criteria tests (Forms A-D) are listed: (1) interpretation; (2) application; (3) analysis of elements; (4) analysis of relationships; (5) production of a plan; (6) derivation of a set of abstract



relations ; (7) judgments in terms of internal evidence ; (8) judgments in terms of external criteria. Form A is a fifty-item, four-alternative, multiple-choice test. The test does not measure the acquisition of information. Rather, it emphasizes research methodology, induction, and deduction. The test was given as a pre-test in September as a post test in January. Since the time interval between test administrations was so great, it was not deemed necessary to use alternative forms for the post-test to guard against criterion contamination. Instructors in all groups saw the examination for the first time when they administered it to their groups as the pre-test. They saw it the second time when they administered the measure as the post test.

A second criterion measure was used to compare achievement in the LE and SE sections. This was the final course examination. It was a 100-items, four-alternative, multiple-choice test. It was based entirely on the textbook and attempted to measure understanding of basic concepts and acquisition of information. Since many different textbooks were used in the SC groups, it was not possible to administer the examination to these sections.

*Measures of student subjective reaction.*—An opinionnaire of twenty-two statements (to be marked true or false) was administered to students in all groups (see Appendix C). Statements dealt with student opinion about level of course organization and difficulty, classroom atmosphere, teacher-student relationship, and the value of the course. The twenty-two statements represent those for which there was 100 per cent. agreement in rating among three judges. The opinionnaire was scored by counting in number of favorable responses (1 to 22). Responses to statement No. 4 were subjected to special analysis using the Chi-square formula (since these data are discontinuous). This statement is : The atmosphere of the class was cold and impersonal. Responses to a twenty-third statement (not included in the total) were similarly analyzed. This statement is : Psychology is based more on *fact* than on feeling. The opinionnaire was administered at the last formal class meeting. Copies were labelled 'confidential' and students were requested not to sign their names. Opinionnaires were identified only by section number.

An Evaluation Scale was used in LE and SE sections (see Appendix D). Using a five-point scale ranging from excellent (5) to poor (1), students were asked to rate particular characteristics of the experimental sections and how much they thought particular characteristics contributed to their learning in the course. There were twelve characteristics listed, so that scores could range from 60 to 12 student ratings for three single characteristics were specially analyzed. These are listed on the Scale as 'Critical Reports,' 'Experimental Project,' and 'Class Size.'

## II.—RESULTS.

The present problem involved estimates of differences between independent population means. In determining the standard error of difference between means, no correlational term was used since Ss. were randomly assigned to the groups and no logical way existed for pairing the Ss. Two-tailed tests were used in determining significance levels.

Using the differences between pre-test and post test results, the mean gains on the *Introductory Psychology Criteria Test* were determined and compared. The results are shown in Tables 1A and 1B. The Z scores of .88 (LE *versus* SE), .51 (SC *versus* LE+SE), and .55 (SE *versus* SC) are *not* significant within .01 or .05 confidence limits.



TABLE 1A  
INTRODUCTORY PSYCHOLOGY CRITERIA TEST—FORM A.  
(Pre-test—Post test for all groups.)

Group type	N	Mean-gain score	S <sup>2</sup>	Z
LE . . . . .	224	3.30	5.09	.88
versus SE . . . . .	163	2.86	4.68	
SE + LE . . . . .	392	2.91	4.91	.51
versus CS . . . . .	290	3.11	5.32	
SE . . . . .	168	2.86	4.68	.55
versus SC . . . . .	290	3.11	5.32	

TABLE 1B  
GROUP MEAN GAINS FROM PRE-TEST TO POST TEST ON THE INTRODUCTORY PSYCHOLOGY CRITERIA TEST.

LE	SE	SC
4.55	4.50	6.72
2.56	2.97	4.52
	2.91	4.10
	2.49	2.78
	2.20	2.65
	1.79	2.46
		2.24
		1.63
		1.97
		1.12

The mean scores for the LE and SE groups on the *Final Examination* were compared. The results are shown in Table 2. A Z score of .42 is not significant within the conventional confidence limits.

TABLE 2  
FINAL EXAMINATION.  
(Experimental Groups Only.)

Group type	N	M	Z
LE . . . . .	227	58.05	.42
versus SE . . . . .	161	59.36	

The analysis of the results of the *Evaluation Scale* are shown in Table 3 (A-D). This Scale was administered to the experimental groups (LE and SE) only.\* Means of 39.34 and 40.41 yield a Z score of 1.44 which is not significant. For the three single characteristics (i.e., the critical reports, the experimental project, and class size) there are statistically significant differences which favor the SE groups.  $Z=2.096$  ( $p < .04$ ) for the critical reports.  $Z=2.31$  ( $p < .02$ ) for the experimental project.  $Z=10.07$  ( $p < .001$ ) for class size.

TABLE 3

## EVALUATION SCALE

## A.—Comparison of Mean Scores for Entire Scale.

Ground type	N	M	S <sup>2</sup>	Z	
LE.....	195	39.34	6.04	1.44	$p < .15$
versus .....					
SE .....	120	40.41	6.37		

NOTE.—The higher the score the more the ratings are favorable.

## B.—Comparison of Mean Ratings for 'Critical Reports.'

Group type	N	M	S <sup>2</sup>	Z	
LE.....	195	3.13	1.19	2.096	$p < .04$
versus .....					
SE .....	120	3.39	1.10		

## C.—Comparison of Mean Ratings for 'Experimental Project.'

Group type	N	M	S <sup>2</sup>	Z	
LE.....	195	3.61	1.30	2.31	$p < .02$
versus .....					
SE .....	120	3.92	1.35		

## D.—Comparison of Mean Ratings for 'Class Size.'

Group type	N	M	S <sup>2</sup>	Z	
LE.....	195	2.61	1.57	10.07	$p < .001$
versus .....					
SE .....	120	3.92	.91		

\* In one SE group the instructor failed to administer both the Evaluation Scale and the Opinionnaire. The results of only five of the six SE sections are, therefore, reported here.



Analysis of the data from the Opinionnaire appears in Table 4A and 4B. For Z scores of .98 (LE *versus* SE), .19 (SC *versus* LE=2E) and .83 (SE *versus* SC), *p*'s are not within conventional confidence limits. For statement No. 4 (Classroom atmosphere) there are no significant differences. However, for statement No. 23 (Psychology as fact or feeling) at statistically significant levels, the experimental sections reported that psychology was based more on fact than feeling.

TABLE 4  
OPINIONNAIRE  
(All groups)

A.—Comparison of Mean Scores for Entire Opinionnaire.

Group type	N	M	S <sup>2</sup>	Z
LE.....	189	15.70	13.04	.16
<i>versus</i> SE .....	121	16.09	11.15	
LE+SE ....	310	15.86	12.66	.19
<i>versus</i> SC .....	316	15.80	9.31	
SE .....	121	16.09	11.15	.83
<i>versus</i> SC .....	316	15.80	9.81	

B.—Statements Nos. 4 and 23.

	df	$\chi^2$
Statement ..4 ....	2	5.15
Statement ..23 ....	2	8.77*

\**p* < .01 — .02

III.—DISCUSSION.

In the present experiment, we hypothesized that class size would be a significant variable in determining the achievement of students in large (LE) as compared with small groups (SE and SC) and for students in co-ordinated (LE and SE) as compared with unco-ordinated (SC) groups. The results of both criteria measures of achievement are not consistent with this hypothesis. The question of the validity of our measures must be raised. The final Examination had high content validity for the LE and SE groups since it was based almost exclusively on the textbooks which Ss were asked to read from cover to cover. The validity of the Introductory Psychology Criteria Test is open to question on two counts. First, the small mean gain for all groups combined was about three points. With seventeen weeks of instruction, one might expect somewhat larger gains. However, at the University of Michigan, where the test was constructed, the mean gain for students was three points higher than ours.

Secondly, the test may not be a valid measure of all the behavior changes which the traditional introductory psychology course attempts to produce. However, before dismissing the criteria test, we should remember that no practical test exists or will exist to assess all the behavioral and attitudinal changes which heterogeneous psychology staffs entertain for a single introductory course. This investigation can only indicate the comparable degrees of success of various classroom conditions in producing the behavior changes which the test attempted to measure.

We also hypothesized that students would significantly prefer one of the class sizes and either co-ordinated or unco-ordinated sections. The results of the *Opinionnaire* do not support this hypothesis. No convincing evidence was found that students prefer small classes to large classes, nor do they significantly prefer either co-ordinated or unco-ordinated sections. They do not report that they find large classes necessarily 'cold and impersonal' nor, conversely, small classes 'warm and personal.' Neither class size nor co-ordination of instruction are apparently the crucial variables in determining students' preferences in the context investigated.

The *Evaluation Scale*, used only in the SE and LE groups, indicates a trend which partially favors the small groups. There is, however, no significant difference between the means of the SE and LE groups. For the critical reports, a major assignment and for the experimental project, the major assignment, the differences clearly favor the SE sections. However, for both assignments the mean ratings in the LE groups were between 3 and 4 (on a five-point scale), indicating the somewhat favorable value the students assigned to these projects. This difference may be due to the smaller instructor-student ratio in the SE sections enabling students more easily to obtain help with these assignments. The teaching assistant-student ratio for the large sections was about 75 to 1, whereas the average instructor-student ratio in the SE sections was 28 to 1. The most significant difference to appear was in the responses to the item 'class size,'  $Z=10.07$ , which is significant at the .001 level of confidence. This indicates that when asked directly about the advantages for learning of small or large classes, students consistently report that the greater benefit lies with the small class size. This suggests that students frequently report that class size is a handicap to learning when they are asked specifically about this factor, even though their responses to the remainder of the scale or to the *Opinionnaire* are not consistent with this expression of attitude. One can at least suspect stereotypic thinking here on the part of the students. We should recall that, in fact, there was no convincing evidence that class size was a handicap to their achievement.

For an experiment of this duration (over seventeen weeks) and one making only gross comparisons, there are many sources of variability which make difficult an appraisal of the effects of the independent variables and the many influences on the dependent variables. The only generalization we can make is that large class size and co-ordinated instruction are not necessarily obstacles to student achievement as defined by our criteria measures. Nor, in the context of this experiment, do students give a convincing and consistent indication that they prefer small to large classes.

What further explorations are suggested? One tack is to find ways of maximizing the effectiveness of large group instruction by improved content and technique. Increased effectiveness might be obtained through the use of small seminars which are co-ordinated with the large group experiences. Large group instruction can be combined with programmed instruction in order to individualize and enrich learning. Methods for the guidance and confirmation of student



responses in large class settings could be experimentally developed. Another tack may be to develop criterion measures which indicate motivational and attitudinal changes which may be produced by particular instructional methods. Such measures would help us define a full range of objectives for the introductory psychology class and enable us to carry on experimentation in which we compare the effects of various instructional techniques, materials, media, and patterns of teacher-student interaction.

In our continued study of class size it may be well to avoid a rigid dichotomizing of practice into exclusively small or large-group and co-ordinated or unco-ordinated instruction. In the interest of efficient and effective instruction, we may try many patterns of shared instructional responsibility (Trump, 1959). We can consider classes of all sizes, from very large groups to the dyadic groups, described, perhaps somewhat wistfully, by Mark Hopkins.

#### IV.—REFERENCES.

- <sup>1</sup> CARL L. AMUNDSON (1954). The class size bogey is here again. *Calif. J. of Second. Educ.*, 29, 111-113.
- <sup>2</sup> B. BLOOM (Ed.) (1956). *Taxonomy of Educational Objectives*. New York: David McKay. (The book by P. L. Dressel and Associates, *Evaluation in Higher Education*. Boston: Houghton Mifflin Co., 1961, bases much of its discussion on the Bloom *Taxonomy*.)
- <sup>3</sup> CAMMAROSANO, J. R., and SANTOPOLO, F. A. (1958). Teaching Efficiency and class size. *School and Society*, Sept. 27th, 338-341.
- <sup>4</sup> HUDSON, E. (1928). *Class Size at the College Level*. Minneapolis: University of Minnesota Press.
- <sup>5</sup> LOVELL, G. D., and HANER, C. F. (1955). Forced-choice applied to college faculty rating. *Educational Psychological Measurement*, 15, 291-304.
- <sup>6</sup> MACOMBER, F. G., and SIEGEL, L. (1960). *Final Report of the Experimental Study in Instructional Procedures*. Oxford, Ohio: Miami University, 96 pp.
- <sup>7</sup> MCKEACHIE, W. J. (1960). Improvement of instruction. *Rev. of Educ. Res.*, 30, 351-360.
- <sup>8</sup> NELSON, W. B. (1959). An experiment with class size in the teaching of elementary economics. *Educational Record*, 40, 330-341.
- <sup>9</sup> RILEY, J. W., RYAN, B. F., and LIFOWITZ, M. (1950). *The Student Looks at His Teacher*. New Brunswick, New Jersey: Rutgers University Press.
- <sup>10</sup> ROHRER, J. H. (1957). Large and small sections in college classes. *J. of Higher Educ.*, 28, 275-280.
- <sup>11</sup> TRUMP, J. LLOYD (1959). *Images of the Future: A New Approach to the Secondary School*. Washington D.C.: National Educational Association.

(Manuscript received 24th June, 1963.)

#### V.—APPENDIX C.

OPINIONNAIRE\*

Copy No.

Psychology 10.1  
Fall, 1962

CONFIDENTIAL—DO NOT SIGN YOUR NAME

Your Section Number  
Mark each statement "T" for true and "F" for false or disagree. Be sure to mark each statement.

1. The organization of the course was more effective than most.
2. Difficult concepts were not explained in class.

\* These statements were selected (and modified) from an opinionnaire which appears in D. M. Johnson, *Teacher's Manual to Accompany Psychology: a Problem-Solving Approach*, New York: Harper and Brothers, pp. 91-92.

# Class Size and Co-ordinated Instruction

3. The teacher was concerned about the students.
4. The atmosphere of the class was impersonal and cold.
5. After being in this class, I have decided to take another psychology course as an elective.
6. In this class my attention wandered frequently.
7. The course is too difficult for freshmen.
8. The teacher made me wonder about things I had not thought of before.
9. Psychology is so easy that I have had extra time to spend on my other classes.
10. The course has been so interesting that I have spent more time on it than I originally planned.
11. Now that I have studied psychology, I notice that I spend more time analyzing behavior and less time criticizing people.
12. The teacher understands college students and their problems.
13. The teacher seldom smiled.
14. A shy person would be afraid to ask questions in this class.
15. The grading was fair.
16. In class I discovered that psychology is a theoretical science that has no particular value for me.
17. During the term I discovered that psychology can be applied to my major field more than I had realized.
18. Psychology is so hard that I have had to take time from my other courses.
19. I found that I had to study almost every day in order to keep up.
20. The course as a whole is a snap.
21. I found that I spent less time in preparation for tests than I expected.
22. The class was, personal and enjoyable.
23. Psychology is based more on *fact* than on feeling.

## VI.—APPENDIX D.

Psychology 10.1  
Experimental Sections  
Evaluation Scale  
Fall, 1962

CONFIDENTIAL—DO NOT SIGN YOUR NAME

Your section number

Please rate the following characteristics and activities of the course.

Use the following scale.

5	4	3	2	1
excellent	good	average	fair	poor

The ratings should be your own—what *you* think. The basis of your ratings should be how much the particular item contributed to your *learning* in this course.

Critical reports.

Films.

Experimental project.

Textbook (Hilgard, 3rd ed.).

Student Guide.

Quizzes.

Mid-term examination.

System of grading.

Class size.

Class lectures.

Class discussions.

Course outline.



# UNTRAINED AND TRAINED GRADUATE TEACHERS A COMPARISON OF THEIR EXPERIENCES DURING THE PROBATIONARY YEAR

BY MILDRED COLLINS  
(*School of Education., University of Leicester.*)

SUMMARY. The probationary year's teaching experiences of untrained and trained graduate teachers, who left the same university at approximately the same time, were compared. This was done by means of questionnaires returned by the teachers themselves and the Heads of their schools. The untrained graduates' replies were matched for sex and academic qualifications with replies received from the trained graduates. Certain areas of experience of the two groups were then compared. Few concessions seem to have been made by the schools for lack of training of their new graduate teachers and the untrained group proved generally to be the less satisfactory group from the schools' point of view.

## I.—INTRODUCTION.

IN March, 1961, of the graduate teachers below 25 years of age in maintained secondary and primary schools in England and Wales, 38 per cent. of the men and 46 per cent. of the women were untrained (Ministry of Education, 1962), that is they had not taken a year's post-graduate course of professional training, they were without the qualification of a Certificate in Education. But they were 'qualified teachers,' for in England and Wales—but not in Scotland—the possession of a degree automatically confers upon the graduate the status of 'qualified teacher,' which he will lose should he fail his post-graduate teacher-training course.

The increase in the number of posts available in schools and the wish of many students to marry early and settle down—together with certain feelings of uncertainty about the value of post-graduate courses for teaching—have resulted in an increase in the proportion going to teach in schools without training. At the same time, these untrained graduates leave the profession more frequently: in 1957-58, the gross wastage rate for untrained men graduates was 6.2 per cent. compared with 2.2 per cent. for those who had trained, while for women the figures were 13.5 per cent. and 8.0 per cent., respectively.\* The untrained graduates are also less satisfactory in their probationary year: in 1957-58 for example, 10.9 per cent. of them compared with 3.9 per cent. of the trained graduates had to have their probationary year extended.† This probationary year is, of course, the first year of teaching in England and Wales and a satisfactory report on the year's work is required before the appointment of the graduate, or non-graduate for that matter, is finally confirmed by the Ministry of Education. An extension of time is usually given to a teacher who has not reached an acceptable standard so that he may have an opportunity to improve. If he does not, then at this point also the graduate may lose his qualified teacher status.

\* Information supplied by the Ministry of Education. The figures are based on records of April, 1957, and March, 1958, and therefore, do not include teachers who entered and left the profession between those two dates.

† Information supplied by the Ministry of Education.



The probationary year is, therefore, an important year for a new teacher : " It is both the beginning of true professional responsibility and the last phase of initial training " (N.U.T. and A.T.C.D.E., 1962, p. 2), and yet there is virtually no organised information about teachers during this period. Colleges and departments of education have many casual contacts with their former students and appointment boards sometimes hear from dissatisfied untrained graduates wanting to leave teaching for another job, but apart from a follow-up study of forty former college students in their first year of teaching (Pearce, 1959) and a more informal survey carried out by another college (N.U.T. and A.T.C.D.E., 1962), I do not know a single piece of research in England devoted exclusively to the probationary year,\* although one has recently been carried out by Nisbet in Scotland, where the probationary period is two years. Neither do I know of any work, devoted to any year at all, which includes untrained graduates. They have been left severely alone. There have been, however, a few follow-up studies of trained graduate teachers (Nisbet, 1954) (Collins, 1959) and teachers from training colleges (Tudhope, 1942) (Rudd and Wiseman, 1962). I propose, therefore, in this enquiry to concentrate attention on untrained graduate teachers, to compare the probationary year's experiences of a group of these teachers with a group of trained graduate teachers.

## II.—THE PLAN OF THE ENQUIRY.

### (1) *Subjects.*

This investigation was in the nature of a pilot enquiry, and the numbers concerned were relatively few, hence it seemed necessary to stabilise some of the variables by taking the untrained graduates from one university and comparing them with a matched group of trained graduates leaving the same university at the same time, or, in a few cases, the previous or following session. The proportion of students entering teaching varies from university to university (B.F.U.W., 1962) and probably also the attitudes and academic level of the students concerned. To deal with one university only minimises these factors.

Among the men graduating from this particular university in 1960 and 1961, 19 per cent. and 21·7 per cent., respectively, did not train ; for women the proportions were 45 per cent. and 40 per cent. It was from these untrained teachers, together with twenty-two who left in 1959 (in order to increase the number of subjects), that the information to be discussed in this article was obtained. There were in the group sixty-three women and twenty-nine men : this represents a much higher proportion of women than in the Department of Education of the same university. Here the ratio was 105 women to 183 men for the sessions ending in 1960 and 1961. Similarly, there was a higher proportion of scientists among the untrained, 36 per cent. compared with 26 per cent. in the Department.

The enquiry about the probationary year's experiences of both the untrained and the trained graduates was carried out by means of two questionnaires, one to the former students, which was answered anonymously, and the other to the Heads of the probationers' schools. The latter required an assessment of the teaching ability of the probationer with comments on her† strengths or weaknesses : this was not anonymous. The questionnaire to the former students was much longer and covered such subjects as the introduction to the school,

\* Since this article was written the training college research group of the Birmingham Institute of Education has published a summary of an enquiry in the *Times Educational Supplement*, Nov. 15, 1963, p. 708.

† As the majority of the untrained graduates were women, the feminine pronoun will be used, but this must be taken to include the men.



forms taught, time spent on preparation, etc., difficulties and satisfactions found in teaching and some personal details as to reading, living arrangements and so on. It was, of course, impossible to match the teachers' questionnaires to the Heads' short version, and this was unfortunate—but intentional.

The response from the untrained probationers was 55 per cent. for the men and 66 per cent. for the women, with the science students, particularly among the women, somewhat 'over-represented.' There were sixteen replies from the men and forty-two from the women, making fifty-eight in all. The response from the Heads of schools was much better, although sufficient information about schools, for an enquiry to be made, was only available in seventy cases: sixty-three replies were received, a 90 per cent. response, and they concerned seventeen men and forty-four women. Slightly more were returned from grammar schools and slightly fewer from primary schools than were received from the probationers, but the differences were negligible. The academic standard of the probationers as stated on their questionnaires, and of those reported on by the Heads, was virtually the same. The probationers' questionnaires and those from the Heads can certainly be taken as referring to the same group of probationers.

As has already been noted the comparison was done by means of a 'matched group' both for the Heads' and the teachers' replies. The matched group was selected from among the questionnaires returned by graduates (and their Heads) who left the Department of Education in 1960 and 1961. These questionnaires represented a 77 per cent. response from the trained graduates and an 84 per cent. response from the Heads. The matching of the group (or groups) was done for sex, degree standard and subject and there were only six minor discrepancies. Academic achievement has been shown to be important in England during training (Warburton, Butcher and Forrest, 1963, Collins, 1958) and in Scotland for professional advancement afterwards (Nisbet, 1954). Hence, it seemed necessary to take it into consideration in this way. We have, then, two pairs of groups of questionnaires on which to base an analysis, one pair made up of replies from untrained and trained graduate teachers, and the other from replies received from the Heads of the schools of the untrained and trained graduates. In each case, the trained group acted as a 'control.'

## (2) *The Points of Enquiry.*

The first point of enquiry was to find out whether the untrained and trained graduates obtained posts in similar types of schools and, once in those schools, whether the conditions under which they worked were the same. Then, in the light of this information—and within the limits set by the data available—to test the following hypotheses by applying the chi-squared test for significance to any differences found in the replies on the questionnaires from the different groups.

- (a) The untrained teachers will be less efficient than the trained.
- (b) They will view the school situation from a different standpoint.
- (c) They will be less committed to their profession.

## III.—THE RESULTS OF THE ENQUIRY.

### *The Schools.*

Table 1 shows the types of schools in which the untrained and the trained graduates obtained their first teaching posts. While the grammar school was the dominant type in both groups, significantly more of the untrained graduates

TABLE 1

TYPES OF SCHOOLS\* IN WHICH UNTRAINED AND TRAINED GRADUATES TAUGHT, BASED ON HEAD'S REPLIES.  
(Figures in parentheses from Probationers' Replies.)

Groups	Grammar	Tech. Coll.	Compre- hensive	Sec. Mod.	Primary	Total Graduates
Untrained ..	31 (27)	0 (0)	3 (2)	18 (16)	11 (13)	63 (58)
Control ....	39 (36)	3 (4)	9 (8)	10 (6)	2 (4)	63 (58)

Heads' Replies :  $X^2=13.18$ . d.f.=3 (grammar and technical together)  $P=0.005$ .

took posts in secondary modern and primary schools, although they had the same academic qualifications as the trained graduates. There was, however, virtually no difference between the groups in the choice of single sexed or co-educational schools: eighteen of the groups together took jobs in boys' schools and the rest were equally divided between girls' and co-educational. The condition of the school buildings, as indicated by the staff room accommodation, appeared to be poorer for the untrained teachers, although the difference was not quite significant at the  $P0.5$  level.

Preparation for a post in a school is often obtained by discussions at the school before actually going to teach there and this did occur in the majority of cases, although with sixteen of the untrained and thirteen of the trained this was not so; neither did eighteen untrained graduates and thirteen trained graduates receive their teaching timetables before the beginning of term.

TABLE 2

ASSISTANCE RECEIVED BY PROBATIONERS.

	From colleagues in connection with :					Visits from Inspectors
	Syllabuses	Lesson Prep.	Exam. Papers	Marking	Badly Behaved Forms	
Untrained ..	23	10	29	21	24	35
Control ....	17	16	36	29	18	22

\* 'School' actually included a few technical colleges as can be seen from looking at Tables 1 and 3, but in the evidence presented in this article, the colleges have not been generally distinguished from schools, as the probationers in them were all teaching some 'O' and/or 'A' level work.

The classification of schools themselves also provided some difficulties. In addition to the types identified by the titles, in the schools from which Heads' replies were received, for untrained graduates 'Grammar' included two technical grammar, two technical high, and two Leicestershire-plan grammar schools: 'Comprehensive' included one private college; 'Secondary Modern' included one intermediate and one Leicestershire-plan high school; 'Primary' included one school for handicapped children. For trained graduates, 'Grammar' included two technical/grammar, one technical high and one Leicestershire-plan grammar school; 'Comprehensive' included one bilateral and one 'selective entry' secondary school; 'Secondary Modern' included two Leicestershire-plan high schools. Similar variations are masked by the classification of the probationers' replies.



The new teacher can also expect to get some assistance from his colleagues and the questionnaire to the graduates included a list of possible areas in which they might have needed and received assistance. Table 2 gives the number of positive replies under the stated headings. None of the differences is large between the two groups, but it is interesting to see that, as with the preparation before beginning the term, it is the trained group which, if anything, has the edge on the other with regard to assistance given. It is true that the untrained graduates had significantly more visits from H.M.Is. (or local Inspectors) but these visits were not necessarily looked upon as an occasion for assistance—and, indeed, half of the graduates did not receive them at all. As one untrained graduate put it: "I received an inspector's visit only as a formality . . . My impression was that one should keep one's fingers crossed not to be 'caught out' during probation."

The work which the two groups was expected to do in schools, showed some variation. While just over half in each grammar school group was concerned with fifth form teaching, it was the untrained group which had the greater responsibility for 'A' level work, two-thirds of the untrained and one-half of the trained graduates in grammar schools were teaching to this standard. The untrained group did, however, have a slightly easier time with regard to form responsibilities in grammar schools (only three probationers were without a form in other schools): one-half of the untrained compared with one-third of the trained were not in charge of a form. The untrained also had more opportunities to observe in different types of schools their colleagues teaching, not that this was a common occurrence, but twelve of the untrained compared with one of the trained had been given, and taken, this opportunity.

These differences are, however, all small and hence unreliable. The general picture which emerges of the conditions under which the two groups of probationers worked, the trained and the untrained, is of essential similarity. The status of 'qualified teacher' given by a degree was taken seriously by the schools: the two groups appear to have been treated much alike. However, one must remember that the two groups were not alike in their tendency to choose primary and secondary modern schools. It is against this difference in type of school but similarity of experience that the hypotheses put forward in this enquiry were tested.

### *Testing the Hypotheses.*

- (1) *The untrained teacher will be less efficient than the trained.*

The 'efficiency' of a teacher can be estimated in various ways (Doman and Tiedeman, 1950) and from various viewpoints of excellence (Robertson, 1957). In this enquiry I am going to take two criteria, first an assessment of the graduate as a teacher by the Head of her school in terms stated on the questionnaire, and secondly, the teacher's record of attendance at school during probationary year. Admittedly, the latter will provide only tangential evidence.

The Heads of the schools were asked to say whether the probationers appeared to him (or her) (a) 'to do well', (b) 'to make a reasonable start and was about average,' and (c) 'to achieve rather less than is usual in the first year': the Heads were also invited, if they wished, to comment on particular assets or weaknesses and to say if the probationary year had not been completed satisfactorily. The distribution of the assessments are shown in Table 3.



TABLE 3

ASSESSMENT OF GRADUATES BY HEADS OF SCHOOLS.

	In All Schools				In Grammar Schools *			In Other Schools		
	Did really well	About aver- age	Less well	Total	R.W.	A.	L.W.	R.W.	A.	L.W.
Untrained	13	35	15	63	9	17	5	4	18	10
Trained ..	30	24	9	63	19	19	4	11	5	5
Total ..	43	59	24	126	28	36	9	15	23	15

In All Schools  $X^2=10.28$ , d.f.=2,  $P=0.01$ .In Grammar Schools  $X^2=2.47$ , d.f.=2,  $P=0.30$ .In Other Schools  $X^2=9.8$ , d.f.=2,  $P=0.01$ .

\* In this classification three technical colleges are put with the grammar schools, and it must be remembered that for the purposes of this article, the term 'grammar school' also includes a few technical/grammar, technical high schools, etc. (see note at bottom of page 3).

When the two groups are considered as a whole, it is clear that the Heads of these schools thought that the trained graduates had done better than the untrained, but when the groups are divided into those who went into grammar schools and those who did not, it is equally clear that the marked difference in achievement—and efficiency in this sense—was due to the graduates who did not go into grammar schools. These probationers, with or without training tended to find—for reasons within themselves or within their schools—the secondary modern, the primary and the comprehensive schools 'tougher', 'going', but it was the untrained who, according to the Heads, found the 'going' hardest. The comments of the Heads underlined this difference in efficiency of the two groups: teaching technique and discipline were remarked on favourably fifteen times for the untrained and twenty-six for the trained, and unfavourably thirty-six times for the untrained and twenty-one times for the trained. Three of the latter did not complete their probationary year satisfactorily, but this applied to seven of the untrained.

No teacher can function efficiently in a school if she is absent for any length of time, not only because of the break in continuity of teaching, but also because of the more prolonged physical and mental malaise which often accompanies it (Hopkins and Malleson and Sarnoff, 1958, and Collins, 1960). The probationers were asked whether they managed to keep well during the probationary year and, if not, how long they were absent from school and for what reason. The trained graduates had been more free from illness, forty-five compared with thirty-seven of the untrained stated that they had not been away from school, and of those who had been absent, fifteen of the untrained compared with six of the trained had been away for one whole week or more. These differences are not significant at the  $P0.05$  level, but they are sufficiently large to help to support the conclusion derived from the Heads' assessments. It seems that the hypothesis that trained graduates would show greater efficiency is substantiated.



- (2) *The trained and the untrained graduates will view the teaching situation from different points of view.*

The school situation can obviously be defined in many different terms and here the limits will be set by taking three aspects only; school discipline, the preparation and marking of school work and the relevance of undergraduate university experiences to teaching. It is argued that the trained graduates, by virtue of their courses in education and their wider knowledge of schools, should perceive underlying factors determining discipline not obvious to a common sense approach, should mark and prepare their work more rapidly and should see more relevance of non-teaching activities to the teaching situation itself.

To take discipline first, the probationers were asked whether the general discipline in their schools was good, satisfactory or rather poor, and what they thought were the factors determining this. Less than one-fifth of the schools were classified by the teachers as 'rather poor' and there was nothing to choose between the discipline of the schools of the trained and the untrained, at least as seen by them. The comments which they made on 'factors determining' discipline have been grouped under several main headings: they are the personality of the Head, strengths and weaknesses of the staff, traditions and expectations of the school, school organisation such as arrangements for staff meetings, supervision, etc., home background of children, and use of corporal punishment, but none of these remarks, grouped under the headings, showed any clear differences between those who had taken an education course and those who had not, although there were altogether 114 comments from the former and 142 from the latter. The only difference in viewpoint which was hinted at was a tendency among the untrained teachers to comment more frequently in terms of the importance of one person, 'a respected Head,' 'a weak Deputy,' 'a good Head of Department': they made forty-five remarks of this kind compared with thirty-three from the other group. Perhaps not unconnected with this was the fact that seven of the untrained teachers mentioned corporal punishment whereas none of the control group did so. The training of the latter did not, however, in the terms of their answers to questions on discipline, distinguish them in any reliable way from their untrained contemporaries, and in fact, what small differences there appeared to be might be explicable in terms of the differences in types of schools in which the trained and the untrained found themselves.

There were two questions on the relevance of their university experiences to teaching: one asked which part of their degree studies proved most helpful in teaching, and the other asked if any activities in university student societies, etc., helped to give confidence or to arouse interest and general awareness in teaching. Rather fewer of the untrained graduates found their studies helpful, but of the sixteen who had studied psychology, twelve said that it proved useful compared with only two of the nine trained graduates with similar qualifications. However, these are not significant differences. When we turn to the relationship of participation in student societies, the position is different, here the evidence is much more definite: twenty-nine untrained graduates, compared with ten trained could see no relationship between these activities and teaching, and this difference is significant at the  $P 0.025$  level and is not due to the untrained graduates having been less active during their degree course. In these particular terms, the untrained graduates did see the teaching situation differently.

Were the untrained graduates at a disadvantage when it came to preparing and marking school work? Did they view it as a greater problem? Did it take



them longer? On average it did not: the mean hours per week which the untrained teachers estimated they spent on marking and preparation in grammar schools and in schools of other kinds were 5.8, 6.0, 5.4, and 5.4 (in the order given) while for trained teachers the figures were 7.7, 7.5, 6.0, and 6.6. This looks as if instead of being slower, the untrained graduates were quicker. Or possibly that, though not quicker, they saw less of what was entailed and were less prepared to give time to their work. Actually, however, in spite of these differences being consistently in one direction, they are not statistically significant and this is because the most striking thing about the stated times of work was their very wide individual variation—from 1 hour to 10 hours per week for preparation by the untrained in grammar schools and from 30 minutes to 20 hours for the trained teachers similarly engaged, from 2 hours to 15 hours for the marking by untrained teachers in non-grammar schools and so on. Nevertheless, the hypothesis that the untrained graduates would spend longer on their work is not substantiated and there is a suggestion that they tend to spend less. Taking the hypothesis as a whole, that the untrained graduate would view teaching differently from the trained, it seems that this has only been clearly supported in one field, the relevance of undergraduate participation in student societies to teaching in schools, and that in the other two areas the results were ambiguous.

(3) *The untrained teachers will be less committed to their profession.*

The starkest test of this is whether the teacher remained in teaching beyond the probationary year, whether they were committed to a second year. The answer is unequivocal: according to the Heads, no less than sixteen of the sixty-three untrained teachers had left the profession by the end of the first year; this only applied to three of the trained. This difference is significant at the  $P 0.005$  and compares interestingly with that for the teachers who left to go to other schools: in this case the figure for both groups was the same, namely ten probationers.

Another way of showing identification with one's job is by belonging to a professional association and by reading professional or semi-professional journals. The teachers were asked questions about reading and affiliation to associations during the probationary year and in both cases a significantly larger number of the untrained belonged to no professional organisation such as the N.U.T., A.A.M., etc. (twenty-three compared with ten of the trained), and read no professional journal such as the *Times Educational Supplement*, *Primary Education*, *The New Era*, *The New Scientist*, etc. (twenty-three untrained compared with seven trained). It seems clear that on both counts, continuation in the profession and ancillary professional activities, the untrained graduate teachers tend to be less committed to their profession.

#### IV.—DISCUSSION AND CONCLUSIONS.

In terms of the answers given on the teachers' and Heads' questionnaires, the differences between the untrained and the trained graduates show the former to have been the less satisfactory group—from the schools' point of view—during their probationary year. By and large, it seems fair to say that teaching for the untrained graduates was an occupation to be taken more casually, less seriously, and a job which made fewer demands. They were more prepared to take posts in schools with which they were less familiar, like secondary modern and primary schools. Once in them, they at least did not spend longer on their work



than did the trained graduates, although virtually they had no concessions for their lack of training. They also viewed teaching in narrower terms, they read fewer professional journals, they saw less relation between the arts of school teaching and those developed by activities in student societies. They did not so closely identify themselves with their profession; two-fifths belonged to no professional association, one-quarter left teaching at the end of the first year. This lack of commitment is probably associated to some extent with the higher marriage rate among the untrained graduates in spite of their being, except in the case of the men of 1961, one year on average, younger: sixteen of the untrained compared with eight of the trained were married by the end of the probationary year. There is, however, in this enquiry, no evidence of the extent to which the lack of commitment could be related to the deliberate use of teaching as a stopgap occupation or to the encounter of unexpected difficulties. These certainly occurred, as could be seen from the comments on the questionnaires. On the face of it, it is not in the least surprising to find that the Heads of schools judged the untrained group to have achieved less during the probationary year.

It is very tempting to put down this superiority of the trained teachers to their post-graduate Education Department Course, but in fact, while this may well be true, it does not necessarily follow. It will be remembered that the two groups were matched for academic subjects and degree standard but no cognizance was taken of personality characteristics—and these have been shown to be important during training (Warburton, 1963) and hence, very likely during teaching proper (Collins, 1959). Are there personality differences between those who want to train and those who do not? Does training attract the more conscientious, the more interested and possibly the less sexually attractive? If it does, then the results of this enquiry might be explicable in different terms and not, or to a less extent, in terms of training. Or, on the other hand, do Education Departments attract a high proportion of the shy and the nervous, those afraid to begin teaching, or those unwilling to accept full adult responsibility and who wish to cling to the shelter of university life? If this is so, and if these students are accepted, for whatever reasons, then the influence of training may be greater than it appears to be here. Obviously, we need to know far more about the students' and Departments' choices, as well as the training which they do or do not receive.

Complacency is unwarranted with regard to the Department Course. One in seven here had done 'less well than is usual' in the probationary year, three had had their probationary period extended, seven had read no professional periodical during the whole year, and their comments on school discipline did not distinguish them reliably from their untrained contemporaries. It would be useful to know if similar results follow from courses in other Education Departments.

For it is clear from the replies of the graduates in this enquiry that the training took place in the Department of Education, if it took place at all. Although a certain amount of help was given by colleagues in lesson and syllabus preparation, marking, and difficulties with badly behaved forms, this can hardly be seen in terms of training, for it is interesting to notice from Table 2 that more help was given in assessing what should have been taught, i.e., in setting examination papers and marking, than in helping the beginner to be sure she taught well, i.e., in lesson preparation. Generally speaking, the untrained graduates were expected to do the same work as the trained and their guidance, apart from the twelve who visited infrequently their colleagues' lessons, was not

markedly different and, certainly did not appear to be greater. Rather one might say that just as the untrained graduates tended to take their work less seriously so did the schools treat them less seriously too. There was little evidence from either group, for the view, that 'probation is an essential part of training' (N.U.T. and A.T.C.D.E., 1962, p. 2).

Is this state of affairs likely to be generally true? Extreme caution, indeed, is required in extrapolating from such a small sample, with conditions varying from university to university (B.F.U.W., 1962), but these results suggest that an untrained graduate going into a maintained school, particularly a non-grammar school, requires more help than he (or she) is at the moment given, or—and perhaps this is crucially important—even knows that he needs.

ACKNOWLEDGEMENTS.—Obviously, this article could not have been written without the help of many people, and I should like to thank the graduate teachers, the Heads of schools, the Ministry of Education, and my colleagues of the University of Leicester for their kind assistance.

#### V.—REFERENCES.

- B.F.U.W. (1962). *Recruitment of Women Graduates to the Teaching Profession*. Report of a Working Party of the British Federation of University Women. Unpublished.
- COLLINS, M. (1958). *Some Correlates of Inadequate Teaching Ability*. Ph.D. Thesis, University of London Library.
- COLLINS, M. (1959). A follow-up study of some former graduate student teachers. *Brit. J. Educ. Psychol.*, 29, 187-197.
- COLLINS, M. (1960). Health, family and teaching. *Universities Quarterly*, 14, 57-66.
- DOMAS, S. J., and TIEDEMANN, D. V. (1950). Teacher competence: an annotated bibliography. *J. Exp. Educ.*, 19, 101-218.
- HOPKINS, J., MALLESON, N., and SARNOFF, I. (1958). Success and failure among university students. *Brit. J. Educ. Psychol.*, 28, 25-36.
- MINISTRY OF EDUCATION (1962). *Statistics of Education Part II*. H.M.S.O.
- NISBET, J. D. (1954). A twenty-year follow-up of intelligence test scores. *The Advancement of Science*, II, 13-16.
- N.U.T. and A.T.C.D.E. (1962). *Teachers in their First Posts*. *The Probationary Year*.
- PEARCE, W. M. (1959). A follow-up study of college students. *Education for Teaching*, 48, 41-48.
- ROBERTSON, J. D. (1957). The attributes of successful student teachers. *Brit. J. Educ. Psychol.*, 27, 115-126.
- RUDD, W. G. A., and WISEMAN, S. (1962). Sources of dissatisfaction among a group of teachers. *Brit. J. Educ. Psychol.*, 32, 275-291.
- TUDHOPE, W. B. (1942). A study of training college final teaching marks as a criterion of future success in the teaching profession. *Brit. J. Educ. Psychol.*, 12, 167-171.

(Manuscript received 26th June, 1963.)



# INTELLIGENCE AND THE IMPROVEMENT IN A GROSS MOTOR SKILL AFTER MENTAL PRACTICE

By K. B. START

(*University of Western Australia*)

**SUMMARY.** Forty-four male teachers' college students mentally practised a gross motor skill with which they had previously been unfamiliar. At the end of the six-day practice period, the subjects' ability to physically perform the movement was assessed and was compared with the scores the subjects had obtained on a test of general intelligence. An insignificant negative correlation was obtained.

In breakdown of the relation between these two variables, it appeared that the highly skilful subjects, as assessed by their physical performance at the test, tended to have a significant negative correlation between their skill scores and the intelligence marks. This feature was noted in another study in the literature.

## I.—INTRODUCTION.

HARRISON (1960), in her survey article, noted that if a physical skill is mentally practised or rehearsed, a significant improvement in subsequent physical performance frequently occurs. Despite the fact that controlled experimentation in the field is a recent feature, 'mental practice' has long been used by the top-level performers in most skills. The gymnast frequently 'thinks' his way through the movement he is about to perform and often his coach will advise him to imagine himself doing the movement, and even to try to 'feel' the movement. Similarly, whilst waiting on the highboard, a diver frequently may be seen mimicking the movements he hopes to perform during his imminent dive.

Since the work of Perry (1939), some fifteen studies have been noted in the pertinent literature and these have been focussed on experimentally demonstrating and statistically examining the improvement that 'mental practice' can produce. In the past two or three years, interest has begun to focus on the basic individual traits which enable an individual to learn by this method and the present study investigates the possible relationship that might exist between general intelligence and the learning of a gross motor skill by mental practice alone.

## II.—THE GROSS MOTOR SKILL SELECTED AS THE CRITERION MOVEMENT.

The motor skill of the single leg upstart was used in this study for three main reasons. It was desired to test a skill of which the subjects had no previous physical experience. This would remove the problem of varied prior experience and might tend to limit the effects to that of innate physical ability (if such a thing exists) and ability to learn by 'mental practice.' Secondly, the skill required specific equipment, access to which could be controlled, and during the experimental period, denied the subjects. This would limit the practice possible during the experimental period to the mental practice which was the variable to be examined. Finally, as mental practice had been shown to be effective mainly in simple physical skills, such as throwing, it was hoped to test its efficiency in a gross skill involving movement of the whole body.\*

\* Pilot studies had shown that this skill could, in fact, be acquired within the practice period envisaged.

## III.—METHOD.

Forty-four first year teachers' college males between the age of  $17\frac{1}{2}$  and  $24\frac{1}{2}$  years (mean = 18.8 years) were given five minutes' mental rehearsal on each of six days.\* On the seventh day they were asked to perform the skill, a single-leg upstart on the Olympic High Bar and their performances were rated by a panel of three experienced judges. The scores obtained by the students on the A.B.40 Test, developed by the Australian Council for Education Research, were obtained from the college records and were taken to be an indication of the general intelligence of the subjects.

## IV.—RESULTS.

The ratings of the four judges were compared by Pearson's  $r$  and the inter-correlations were between .92 and .96. The mean rating was tested for reliability by the method suggested by Ebel (1951) and was shown to have a reliability of .96. It was felt that this result permitted the use of the ratings as a criterion measure. The mean ratings were then T Scaled and compared with the I.Qs., obtained from the A.C.E.R. Test. The correlation so obtained was  $-.084$ .

This result agrees with the many studies which, over the past thirty years, have shown that intelligence has low to trivial correlations with various estimates of physical performance, and with the findings of such studies in motor learning as Ruch (1925), Brace (1927, 1941, 1946), Atkinson (1929), Goodenough and Brian (1929), Thompson and Witryol (1946). It also agrees with the earlier studies, Start (1960), Whitely (1962), which also investigated the relation between intelligence and the efficiency of learning by mental practice.

## V.—DISCUSSION.

The relation that exists between physique and the intellect or motor skill and academic attainment has provided considerable discussion in the literature.

Ausubel (1958) noted that "although physical and intellectual development, for example, initially influence each other, there is no necessary relation between their respective rates of growth in a particular individual" (p. 108). Thompson (1954) noted that physical growth paralleled intellectual growth in young children and Miles (1954) in discussing intellectually gifted children, found they were superior in thirty-four anthropometric measures and in general physical development, but commented "the weakness of the gifted appeared to be most often in subjects requiring manual co-ordination" (p. 999).

It is interesting to conjecture whether this weakness was relative to their other performances or to the population standards.

Ausubel (1958), in his book on *Child Development*, pointed out that "studies of the relationship between physical status and motor ability on the one hand, and intelligence and intellectual ability on the other hand, show negligible or zero correlation through adolescence and adult life" (p. 108), and suggested that the relatively high positive correlations which are found during infancy occur because "many of the tests of intellectual ability included in infant intelligence scales really measure sensory-motor or neuro-muscular skill" (p. 108).

\* Thursday and Friday of one week and the Monday, Tuesday, Wednesday, Thursday of the next week.



Johnson and Capobianco (1959) compared general physical condition and ability, to the learning ability of mentally deficient children and found a very low relation. Howe (1959), on the other hand, compared the motor skills of mentally retarded and normal children and concluded "The normal children were consistently superior to the mentally retarded on a variety of motor skill tasks" (p. 354). This finding agreed with those of Francis and Rarick (1957), Fait and Kupferer (1954) and Rabin (1957). The positive relationship which has been found at the lower end of the I.Q. scales, Ausubel (1958) explained, as resulting from the low intelligence children "who frequently do not possess sufficient general intelligence to learn even simple motor skills" (p. 109). This was advanced as the real reason for low motor performance scores rather than any real lack in motor co-ordination to perform the movements.

In the field of motor learning, Munn (1954) found that sensori-motor learning in young children "is unrelated to scores on intelligence tests" (p. 398). Similarly, Husband (1931, 1935), in his experiments on maze learning, found there were low correlations between intelligence and trials necessary to learn the maze, errors made in running the maze and the time taken to complete the maze, and noted that in uninstructed groups the relation between the learning of two mazes and intelligence were  $\cdot 15$  and  $\cdot 24$ , but with groups that underwent an instructional programme this relationship improved to  $\cdot 43$  and  $\cdot 63$ , respectively. Thus, teaching appeared to benefit the more intelligent subjects to a greater degree. Ruch (1925) noted that the relation between intelligence and the performance of certain motor tasks increased with the complexity of the tasks and reported correlations of  $\cdot 31$  to  $\cdot 71$  in support of this. The only studies noted which previously had compared intelligence with gain from mental practice were those of Start (1960) and Whitely (1962), both of whom reported insignificant correlations. The present findings of  $-.084$  would, therefore, appear to agree with previous evidence in the general field of motor learning and with the earlier mental practice study in particular, such that DiGiovanna's (1937) conclusion, "There is no definite correlation between intelligence and athletic ability, and intelligence and motor ability in college men" (p. 101), would not be refuted on the evidence of the present study.

Whilst it was evident that, within the group as a whole, intelligence was unrelated to performance, it was noted that when the skill raw-scores were looked at in detail before T Scaling, there was a bimodal distribution and that the high scoring subjects appeared to differ from the less skilful in their relation to intelligence. When the sample was sub-divided into the skilful and the unskilful subjects, and these sub-samples separately correlated with the intelligence scores, it was noted that the negative correlation among the highly skilful was significant at the 1 per cent. level, whereas that among the unskilful was insignificant. The difference between the mean intelligences of the skilful and the unskilful groups was found to be statistically insignificant as was an examination of the data by the biserial technique. Why intelligence should be inversely related to performance in the highly skilful, posed many questions.

There is some experimental evidence (Bull, 1958) to support the view that status through physical performance is sought by those whose academic attainment fails to provide the rewards necessary to an individual's ego. This motivational factor would account for more low I.Q. people appearing in the high performance group, but in this study the insignificant difference between the mean I.Qs. of the high and low skill groups actually favoured the more skilful group. Some doubt has been cast upon this substitution theory by the work of Fulton and Prange (1950) who showed that there was not a significant difference

between the motor learning of socially acceptable and socially unacceptable team-mates. Also in a recent study (Start, 1961) of the relation between the streaming and intelligence of schoolboys and their games performance, it appeared that these factors were uncorrelated. However, whilst this substitution theory might have explained a significant biserial between the groups A and B, it did not explain why the relation of the Criterion Score to I.Q. was so marked in the skilful group in comparison to the insignificant relation found in the unskilful group.

In the literature the only study cited which had a breakdown of the group correlation between intelligence and motor performance was that of DiGiovanna (1937) who sub-divided his sample\* into four sub-groups and within each of these sub-groups the I.Q. scores were correlated with 'athletic ability'. His results were very interesting and are summarised below in Table 1.

TABLE 1

CORRELATION OF GENERAL INTELLIGENCE AND ATHLETIC ABILITY (DATA FROM DIGIOVANNA (1937).)

Athlete Grade	N	Correlations with I.Q.
A	107	— .20†
B	78	— .10
C	63	+ .22
D	42	+ .10
Total	295	+ .08

† Significant at 5 per cent. level.

The curvilinear nature of the relation evidently masked a negative correlation significant at the 5 per cent. level for the capable athletes. The correlation trends found in the present study which had a considerably smaller experimental group, appear to follow those of DiGiovanna's data. The basic factor which determined the changing relation between skill and intelligence, did not appear to be intelligence itself, but rather some other trait not yet ascertained.

In the present study, the performance had been taken as a measure of learning, whilst in DiGiovanna's work, attainment was taken for what it was—a cross-product of learning, experience, motivation and innate ability. In fact, any learning situation involves innate ability and motivation and thus, it is problematical whether the significant correlation noted in both studies between I.Q. and Skill is derived specifically from learning mechanisms, or motivation or natural ability. Since ability to learn and natural ability, are perhaps synonymous, the traits necessary for very successful motor learning, are possibly different from those which tests such as the A.C.E.R., A.B.40, of general intelligence, are designed to assess. If the latter is associated with reasoning, either verbal or mathematical, then it is possible to suggest a situation which might account for the data in this study, and those of DiGiovanna.

\* 295 college men between the ages of 18 and 21, i.e., very similar to the sample used in the recent study.



When the motor learning is relatively inefficient the rationalisation of techniques, which a highly intelligent person might evoke, might assist his learning and possibly his physical performance. However, where the motor learning processes are highly efficient the more remote analytic approach of the highly intelligent may, in fact, delay learning due to the possibilities of alternative methods which might appear contradictory. This postulation of a motor learning mechanism does not imply that the learning theory at its base is any different from that involved in the acquisition of the more intellectual skills. Indeed, it is probably that it will have a similar multifactor basis but that the factors which contribute to it might be considerably different from those which have been shown to be efficient in learning intellectual tasks.

## VI.—REFERENCES.

- ATKINSON, W. R. (1929). Relation of intelligence and of mechanical speeds to the various stages of learning. *J. Exp. Psychol.*, 12.2, 89-112, April.
- AUSUBEL, D. P. (1958). *Theory and Problems of Child Development*. New York: Gruen and Stratton.
- BRACE, D. K. (1927). *Measuring Motor Ability*. New York: A. S. Barnes.
- BRACE, D. K. (1941). Studies in the rate of learning gross bodily skills. *Research Quarterly*, 12.2, 181-185.
- BRACE, D. K. (1946). Studies in the learning of gross bodily motor skills. *Research Quarterly*, 17.4, 242-253.
- BULL, K. B. (1958). Investigations into the relationship between physique, motor capacity and certain temperamental traits. *Brit. J. Educ. Psychol.*, 28, 149-154.
- DiGIOVANNA, V. G. (1937). A comparison of intelligence and athletic ability for athletic men. *Research Quarterly*, 8.3, 96-106.
- EBEL, R. L. (1951). Estimation of the reliability of ratings. *Psychometrika*, 16, 407-424.
- FAIT, H. F., and KUPFERER, H. (1954). A study of two motor achievement tests and its implication for the planning of physical education activities for the mentally retarded. *Am. J. Mental Deficiency*, 60, 729-732.
- FRANCIS, R., and RARICK, G. L. (1957). *Motor Characteristics of the Mentally Retarded*. Madison: University of Wisconsin.
- FULTON, R. E., and PRANGE, E. M. (1950). Motor learning of highly chosen and unchosen team-mates. *Research Quarterly*, 21.2, 126-131.
- GOODENOUGH, F. I., and BRIAN, C. R. (1929). Certain factors underlying the acquisition of a motor skill in pre-school children. *J. Exp. Psychol.*, 12, 127-155.
- HARRISON, V. F. (1960). Mental practice and motor learning. *J.O.N.P.E.R.*, March, p. 75.
- HOWE, C. E. (1959). A comparison of the motor skills of mentally retarded and normal children. *Exceptional Child*, 25-3, 352-354.
- HUSBAND, R. W. (1931). An analysis of methods of human maze learning. *J. Gen. Psychol.* 39, 258-277.
- HUSBAND, R. W. (1935). Influence of knowledge and instruction in maze performance. *J. Gen. Psychol.*, 46, 213-218.
- JOHNSON, G. O., and CAPOBIANCO, R. (1959). Physical condition and its effect on learning in trainable mentally retarded children. *Exceptional Child*, 25-3, 352-354.
- MILES, C. C. (1954). Gifted children. *Manual of Child Psychology* (Ed. L. Carmichael). New York: Wiley, Ed. 2, 319-324.
- MUNN, N. L.: Learning in children. *Manual in Child Psychology* (Ed. L. Carmichael). New York: John Wiley, Ed. 2, 374-458.
- PERRY, H. M. (1939). The relative efficiency of actual and imaginary practice in some selected tests. *Arch. Psychol.*, 34, 5-75.
- RABIN, H. M. (1957). The relationship between age, intelligence and sex to motor proficiency in mental defectives. *Am. J. Mental Deficiency*, 62, 507-516.
- RUCH, G. M. (1925). The influence of the factor of intelligence on the form of the learning curve. *Psychol. Monog.*, 34, 160.
- START, K. B. (1960). Relationship between intelligence and the effect of mental practice on the performance of a motor skill. *Research Quarterly*, 31.4, 644-649.
- START, K. B. (1961). The relationship between the games performance of a grammar-school boy and his intelligence and streaming. *Brit. J. Educ. Psychol.*, 31.2, 208-211.

- START, K. B. : (1962). The influence of subjectively assessed games ability on gain in motor performance after mental practice. *J. Gen. Psychol.*, 169-172.
- THOMPSON, G. C., and WITRYOL, S. L. (1946). Relationship between intelligence and motor learning as measured by a high relief finger maze. *J. Psychol.*, 22, 237-246.
- THOMPSON, H. (1954). In chapter on Physical Growth in *Manual of Child Psychology* (Ed. L. Carmichael). New York : Wiley, Ed. 2, 319-324.
- WHITELY, G. (1962). The effect of mental rehearsal on the acquisition of motor skill. Diploma in Education Dissertation, Manchester University.

(Manuscript received 11th June, 1963.)



## RESEARCH NOTE

### A STATISTICAL RE-ANALYSIS OF GETZELS AND JACKSON'S DATA

R. W. MARSH

(*New Zealand Department of Education*)

Getzels and Jackson (1962) set out to investigate just how great the acknowledged discrepancy is between results from conventional intelligence tests and measures of creative potential derived from other sources. They felt they would find proof that the discrepancy, rather than being a function of the technical inadequacy of I.Q. tests, is determined by abilities of a quality quite different from those reflected in the usual tests of intelligence.

To do this they took two groups of adolescents: one high on creativity and not comparably high on I.Q., the other high on I.Q., and not comparably high on creativity. They proceeded to administer the same battery of psychological tests to these two groups, and then compared the ways in which the group responses differed in both quality and quantity. Their results led them to conclude that, while creativity is slightly related to intelligence, it really constitutes a separate cognitive factor.

For various reasons, their conclusions appear to be misleading. Both Burt (1962) and De Mille and Merrifield (1962) have already discussed some of these reasons at length. The present paper concurs generally with their statements, and gives further detailed reasons why Getzels and Jackson's conclusions cannot be regarded seriously. These objections are of two kinds: the first rests on methodological grounds, the second on the significance and meaning of their correlations.

One major query lies with their design. By selecting their groups in the above manner, the authors confounded their variables. Where confounding has been resorted to, it is necessary to provide a measure for possible interaction effects, and the way in which these groups were selected promotes, rather than precludes, such effects. However, in their analysis of the data, Getzels and Jackson utilised a simple Chi Square technique, which is insensitive to the possible interactions mentioned above.

It is also profitable to look at the material itself, to see to what degree it really supports the conclusions. Page 20 of the book gives a table of intercorrelations between tests of intelligence and the creativity tests. Immediately below, the authors claim that the correlations are relatively low for this kind of data, and so imply, as they later point out, the existence of another factor which they call creativity.

Now close inspection of the data shows that this claim is not really justified by the evidence. There are two features that must be accounted for before we can begin to accept the correlations as being widely representative of the variables that make up the table. First, some provision must be made for the effects of the explicit selection of the I.Q. variable. Secondly, some adjustment must be made for the degree of unreliability of Getzels and Jackson's creativity tests.

For the purposes of the following calculations, the correlations, given separately for boys and girls, on page 20 of the book, have been combined.

Sufficient information is given by the authors to enable Gulliksen's (1950) correction for explicit selection to be applied to the correlations between I.Q. and the other variables. The corrected correlations yielded by this procedure are as follows:

I.Q. with Word Association  
I.Q. with Uses  
I.Q. with Hidden Shapes  
I.Q. with Fables  
I.Q. with Make-up Problems

0.48 (was 0.37).  
0.26 (was 0.17).  
0.42 (was 0.33).  
0.18 (was 0.12).  
0.38 (was 0.31).

I have also calculated the values obtained for the correlations between I.Q. and the five variables listed above after a further adjustment has been made for attenuation (Gulliksen 1950). A reliability coefficient of 0.90 was assumed for variables 1 and 3, and of 0.80 for variables 3, 4 and 5. These are conservative estimates; for the higher the reliability of the test the smaller the correction for attenuation. The results of this procedure were as follows: for Word Association, 0.51 (0.37); for Uses 0.30 (0.17); for Hidden Shapes, 0.46 (0.33); for Fables, 0.28 (0.12); and for Make-up Problems, 0.44 (0.31), with the original figures in parentheses.

The above results show that the relations between the variables concerned are stronger than casual inspection of the table would lead one to believe. Furthermore, these figures now express estimates based on the results of 533 subjects, thus all the correlations exceed significance at the 0.001 level.

Even these corrected figures are not fully representative of the degree of relationship between the test variables. The I.Q. scores are corrected for age while the creativity test results are not. In correlating I.Q. with raw scores from the creativity tests, the authors ignored this age factor and produced a coefficient that is only a partial index of the relationships involved. Thus, these corrected figures are still not as high as they would be if all the relevant factors had been precisely accounted for.

However, even an analysis of the table of uncorrected correlations shows that the relations between the contributing variables differ significantly from what Getzels and Jackson suggest. An application of Tryon's modification of a method suggested by Holzinger and Harman (Fruchter, 1954) for grouping tests produced the following results: cumulatively adding variables, 5, 1, 6 and 2 to variable 3 furnished successive B coefficients of 1.68, 1.50, 1.72 and 1.50, respectively, while clustering 4 with 2 independently gave a B of 1.14. Roughly interpreted, this means that the greatest source of variance in all variables, except 4, was the same. Now if Getzels and Jackson's assumptions were correct, the variable outside the cluster should be 6 and not 4. Moreover, the variable with the highest B (belonging) coefficient in the cluster is 6, which is quite the opposite of what their prediction implies. This evidence is further supported by results from a factor analysis. Listed below are the results of my working of Getzels and Jackson's table of correlations, showing the unrotated centroid loadings.

TABLE 1

Test	Factor I	Factor II	Factor III
1 Word Association .....	0.74	-0.21	0.21
2 Uses .....	0.48	-0.34	0.11
3 Hidden Shapes .....	0.61	0.35	-0.15
4 Fables .....	0.43	-0.31	-0.27
5 Make-up Problems .....	0.67	0.22	-0.18
6 Intelligence Quotient .....	0.52	0.26	0.35

The reader can make his own interpretation of exactly what these results would imply if rotated. One thing does seem certain, however, and that is that any or all of these factors may be named more appropriately by invoking conventional intellectual factors than by styling them as some factor of creativity.

To avoid unnecessary controversy, let us regard the first factor as nameless but, nevertheless, representing the common variance between the variables in the table. The size of the loadings on this factor does appear to show that the relation between the several sources of variance in the table is quite large, contrary to Getzels and Jackson's claim.



A factor analysis by simple summation with reduced communalities produced three significant factors, accounting for 50 per cent. of the variance indicated in the correlation table. Of this, 66 per cent. accounted for by the first factor, i.e., by the general factor common to *both* intelligence tests and tests of creativity. Some of the remainder will be error variance generated by the unreliability of the tests. The difference between the correlations before and after they were corrected for attenuation gives a hint of just how great this source must be.

Another feature that must be considered is the relative heterogeneity of the intellectual domains represented by the tests involved. From their descriptions one can assume that a fair proportion of the creativity tests—all except one in fact—are strongly associated with verbal abilities and skills, while the exceptional test—'Hidden Shapes'—is related to spatial problems and abstract things. The results of the factor analysis confirm this preponderance. Thus, considering the range of intellectual variables in the sources of data, it is not surprising that the three factors do not account for still more variance.

The arguments advanced above and the detailed analysis of the author's correlations would seem to indicate that creativity is determined by different phenomena than those Getzels and Jackson assume. Rather than being almost independent of the general factor of intelligence, this factor is the most constant and conspicuous ingredient. Also, the ordinary types of intelligence test, particularly those used by Getzels and Jackson, are not fully representative of all aspects of intelligence. Despite this, the conventional I.Q. is still the best single criterion for creative potential. This may be improved upon by the use of other tests as well.

## REFERENCES.

- BURT, C. (1962). Creativity and intelligence (Critical Notice). *Brit. J. Educ. Psychol.*, 32, 292-298.
- DE MILLE, R., and MERRIFIELD, P. R. (1962). Creativity and intelligence. *Educ. Psychol. Meas.*, 22, 803-808.
- FRUCHTER, B. (1954). *Introduction to Factor Analysis*. New York and Toronto: Van Nostrand.
- GETZELS, J. W., and JACKSON, P. W. (1962). *Creativity and Intelligence*. London and New York: Wiley.
- GULLIKSEN, H. (1950). *Theory of Mental Tests*. New York: Wiley.

(Manuscript received 17th June, 1963.)

## BOOK REVIEWS

CHALONER, L. (1963). *Feeling and Perception in Young Children*. London : Tavistock Publications, p. 105, 12s. 6d.

This well-written and 'down to earth' little book should prove valuable to many young parents and others in contact with children in their early years. Its great merit is that throughout children are spoken about as *people*—individuals who need as much understanding and tact in our relationships with them as do our other friends.

"Acceptance of a child's individuality *as he is*, gives us the first real contact with him that can make for any helpful subsequent relationships . . ." Mrs. Chaloner has a real power to feel with children and to see their point of view. Her remarks on habit training are the epitome of informed commonsense, that works to the results rather than to some theoretical 'oughts.' She shows, by her advice, though not making the point explicitly, that she recognises that you can teach children only what they are able to learn and that neglect of this principle may sometimes lead them to learn exactly the opposite of what you wish! "It is so easy to forget" she says, "that in contending too fiercely with a child who wants his own way we may be giving him a first-class demonstration of the store we set by getting ours!"

The chapter on Speech and Social Development should prove most helpful to all those responsible for young children, particularly as it stresses the part played by the relationships with the adult in this and the last chapter on 'How Children Think of Death,' deals with a difficult matter with a restrained and dignified use of personal experience which at least reveals the difficulties if it cannot solve them.

It is not easy to perceive the shape of the book : it reads like a series of articles rather than an attempt to survey the whole field. Some chapters for instance, deal exclusively with young children, while others have helpful forward links with later phases of development. There are inevitable omissions : it seems a pity, for instance, that in a book with this title, no mention should be made of children's learning which is inextricably linked with perception and feeling.

The lively use of contemporary research should make many people wish to consult the original papers.

Altogether, this seems a sound, clear and simple account of some of the principles which should guide us in seeking to understand or influence the behaviour of the "strangers in the house."

M. BREARLEY.

DONALDSON, M., and WITHRINGTON, D. (1963). *A Study of Thinking*. London : Tavistock Press, pp. viii+263, 35s.

This is an interesting and thoughtful, though rather slight, book. In effect, it amounts to an analysis of the answers of some forty children, aged 10 to 14, to a few examples of five kinds of logical or reasoning problems, similar to those often used in intelligence tests. In the introductory chapters, the authors point out, as others have before, that intellectual skills develop progressively in childhood, through interaction with environment and not merely through internal maturation ; also that we have no satisfactory way of specifying just which skills to include in our concept of intelligence. By applying a Piaget-type clinical approach to individual children's responses, they hope to throw further light on the processes involved, and to provide a better basis for constructing intelligence test items. Errors are classified under three main headings : (a) Structural, i.e., failure to grasp the principle, or appreciate the relations, involved ; (b) Failure to adhere to the constraints of the problem, e.g., introducing irrelevancies from personal experience ; (c) Executive, i.e., success in (a) and (b) but mistakes in applying them to the data. These are extensively illustrated and sub-classified, and their incidence over the age range tabulated. But it is difficult to see what comes out of it all. The main



positive conclusions that the present writer was able to discern were that children fluctuate in tackling such problems and sometimes make errors where previously successful; that multiple-choice test items often fail to reflect their natural error-tendencies (though this could surely be overcome by basing the 'distractors' on answers to open-ended questions); and that Piaget's theory of formal operations and combinatorial schemata does not cover all the developments in thinking skills that occur around the age of 12. Though there are many acute comments on the views of other writers, it is a pity that no reference is made to Hunt's book, nor to the rather similar study of group test errors published in this *Journal* by Lawrence, in 1957.

P. E. VERNON.

GEORGE, F. H. (1962). *Cognition*. London: Methuen, pp. 309, 32s. 6d.

The main object of this book, according to its writer, is to provide "a survey and re-interpretation of existing knowledge of cognitive processes. It is intended to be something between a textbook and a series of more or less original papers on cognition." The author's use of the generic (and perhaps now unfashionable) term, 'cognition,' is meant to indicate that he is concerned, not simply with learning, but also with perception, thinking, reasoning, remembering and even imagining. What we are offered is, in fact, a book which aims to give to these diverse activities the theoretical unity and conceptual clarity which hitherto they have so notoriously lacked.

In pursuit of this ambitious purpose, Dr. George provides chapters on the nature and definition of cognition; on conditioning, learning and the classical theories of learning; on perception and 'Gestalt,' 'Transactionalist' and philosophical theories of perception; and on memory, thinking, language and ethology. All these chapters expound their subjects in an elementary, and sometimes irritatingly offhand, way, and prove little that a final year Honours Psychology student would not be expected to know. In fact, the suspicion grows that they owe their origin to lectures given to undergraduate students. But it would be unfair to judge the whole book in these terms, for chapters 12, 13 and 14 are clearly intended to illuminate in a more general way our understanding of the empirical findings and piecemeal theories surveyed in the remainder of the book.

The prospect is exciting—but the result disappointing. Essentially, Dr. George urges: (1) that psychologists should try, more so than they have done in the past, to develop a *general* theory of cognition; (2) that, in attempting this, they should make more use of the concepts and methods of disciplines related to psychology, such as, for example, philosophy, physiology, cybernetics, ethology, and linguistics; and (3) that, of these, philosophy and cybernetics are the most relevant and helpful, the first because it can help clarify and make more precise the language we use to describe cognitive processes, and the second because it provides a theoretical model by reference to which cognitive processes may be predicted and explained. To many readers much of this will be acceptable and unexceptional. But, it is in the detailed exemplification of exactly how following these recommended procedures does, in fact, increase our understanding of the problems of cognition that this book remains unconvincing. The reader is left with the curious feeling that Dr. George's theoretical and linguistic considerations and the empirical data they are supposed to illuminate, somehow, hardly make contact. It is, for example, distinctly doubtful whether anyone actually struggling with the explanation and control of cognitive processes, whether in classroom or laboratory, will find his chapters on 'a theory language for cognition' and on cybernetics and computers particularly helpful.

Moreover, the book has other deficiencies. It is, in places, written in too casual a style; some references given in the text either do not appear at all in the bibliography or else are shown there with a different date; and there is an unhappy number of inaccuracies and oversimplifications to be found in the expository chapters. Altogether, this is not an encouraging book.

G. W. PILKINGTON.



KEEHN, J. D. (1962). *The Prediction and Control of Behavior*. London: Constable and Co., Ltd., pp. 268, no price given.

Another introductory text.

(1) It is unusually short, its format is attractive, its typography is beautiful. In all these respects it approaches perfection.

(2) It is unusual also in being explicitly Skinnerian in approach and in providing, therefore, a useful elementary account of operant theory. Where Skinner is less obviously relevant, it is ordinary and rather dull. But—

(3) It is intended for beginning students, and it is of little use to them because (a) it wastes too much time on protest and psychological apologetics; (b) the writing is too elementary for the attempted sophistication of some of the contents; (c) the style is often obscure and confusing; and (d) readers of this *Journal* might well conclude that, if this is psychology, it has little to offer to the teacher (or the social worker, or the nurse, or the physiotherapist, or the doctor, or . . .) in training.

Definitely this is not a hit, and it is not recommended. GEORGE SETH.

MORRIS, R. (1963). *Success and Failure in Learning to Read*. Oldbourne, pp. 176, 17s. 6d.

The theme of the six chapters of this book is "the problem of seeing the acquisition of literacy as a developmental process extending from first beginnings to the achievement of more mature reading habits." Nine approaches to reading are discussed to show the effects that methods of teaching reading during the early stages may have on adolescent and adult reading skills, habits and tastes. The methods are, imitation, memorization, spelling, kinaesthetic, phonetic, sounds in colour, word and sentence, and the non-oral method in which reading is taught from the beginning by directly linking the printed symbols with meanings. A story, rather than a mechanical approach is then recommended because this is more likely to develop the level of maturity in reading which will lead to the intelligent use of reading in everyday life. In contrast, the analytical approach in which concentrated practice is given upon graded material designed to extend vocabulary and build up habits of grasping details, etc., may lead to the unreflective reading of material with indifferent content in later life.

A chapter is devoted to a critical appraisal of tests of reading and of comparisons made between them and scores on intelligence tests in the form of Achievement Quotients. It is shown that comprehension tests may distort teaching if they contain items of a factual and vocabulary type, for teachers may be side-tracked from the encouragement of reflective reading of varied books, an essential if adults are going to seek for the proper meanings of passages in literature. The use of the A.Q., and of the static notion of mental 'capacity' as the measure of potential is then contrasted with the developmental approach to intelligence which sees it as built up step by step and concept by concept as the mind grapples with its environment.

The final chapter deals briefly with the treatment of backwardness in reading. The effects of remedial treatment upon the hard core of genuine reading disability cases is not treated, nor is research in this field mentioned, because the author is chiefly concerned with the rehabilitation of the inadequately taught. The successful remedial teacher will use individual work to develop powers of communication through genuine reading experiences, rather than through a stertyped remedial programme of diagnostic tests, remedial exercises and analytical teaching.

This book represents a genuine attempt to see the teaching of reading as the guiding of children's thinking processes rather than the inculcation of tricks with mechanical reading part-skills. For this reason, it is a valuable corrective to the present torrent of graded schemes and programmed kits which are flooding into school classrooms.

J. E. COLLINS.



OLIVER, R. A. C. (1963). *An Experimental Test in English*. Manchester: Joint Matriculation Board, pp. 56.

This is a valuable report since it records the systematic experimentation undertaken by an examining board in preparation for setting one of its examinations. The experimental test referred to is a Use of English paper set by the Joint Matriculation Board to a sample of 764 candidates in October, 1962, as a precursor to the open paper of March, 1963.

The test represents an interesting compromise. In the year when the American College Entrance Board took one step in our direction by replacing one-third of their objective tests by essay questions (in a one-hour paper entitled 'English Composition Test'), Professor Oliver has taken two steps in the opposite direction by including two hours of objective questions in this three-hour Use of English paper.

The remaining hour is given to two samples of the candidates' continuous writing, and an attempt is made to differentiate between them by providing in the question paper material for the first piece of writing and inviting personal views in the second.

Each of a candidate's two compositions was marked by a different team of two markers. Evidence is given that two markers are better than one and that the change of markers for the second composition further improved the final assessment. It would have been valuable to know how much would have been gained by having a third or fourth marker for each composition.

The differentiated double sample of a candidate's writing is an interesting feature of the experiment and might have been more fully reported on. Correlation between the two sets of marks was no higher than .319 and yet no reference is made to this as an aspect of reliability.

Reliability of the test as a whole is, in fact, considered only in the form of mark/re-mark coefficients for a sample of 148 scripts. From this limited aspect the test is highly reliable (.982), but credit for this must go to the objective questions, in which high mark/re-mark coefficients are to be expected and where some form of test/re-test coefficient would surely be a more realistic indication of reliability.

The objective questions are also shown to be more valid than the compositions, taking teachers' estimates as the criterion of validity. We are not told whether the inclusion of the compositions raised or lowered the validity of the test as a whole, and this omission may perhaps hint at the uneasy basis of the whole compromise. The objective questions are more reliable and more valid and are clearly included for that reason: the compositions seem to be included on quite different grounds: "It is universally agreed that the writing of continuous prose is an essential element in the use of English" (p. 40).

The underlying issue here is the extent of the responsibility of examining boards, and it is an issue on which Professor Oliver and the reviewer take opposite sides. The principal argument against special testing techniques as opposed to job-samples rests, in the reviewer's opinion, upon the fact that external examinations are an integral part of our school system, and examiners and teachers must work in collaboration. To-day, for the first time, it is being recognised that examining boards must accept responsibility for the 'backwash effect'—the effect of examination upon teaching. The pattern of this experimental test, and the explicit terms of its justification (p. 38), represent a set-back at a particularly dangerous point. Few can suppose (and they do not appear to include Professor Oliver) that this examination has been introduced simply as an additional selection device: in which case it represents, in fact, a deliberate attempt to have an effect upon the teaching in schools.

The effect of the test in the form used here will inevitably be the publication and the use in many sixth forms of collections of interlinear tests and multiple-choice vocabulary and comprehension tests. This is not America—ours is, in fact, a unique situation, with its free choice of syllabus for teachers and free enterprise in textbook publishing, and we must face the consequence that wherever there are external examinations they will inevitably call the tune—not in all schools but in the majority.

JAMES BRITTON.



PENROSE, L. S. (1963). *The Biology of Mental Defect* (2nd Edit.). London : Sidgwick and Jackson, pp. xxiv + 374, 42s.

Mental defect is a subject which generates excessive emotion and fosters a great deal of woolly reasoning and speculation, often presented as hard fact. The 1949 edition of Penrose's *Biology of Mental Defect*, and its 1953 revision, have stood out in providing a rigorous analysis of some of the problems, coupled with a warm approach to the defectives themselves. Much of the earlier editions was sufficiently definitive to be almost timeless. However, human genetics, which forms one of the main pivots of Penrose's approach, has recently advanced rapidly. For instance, human cytogenetics became an effective study only after 1956. The present volume is a completely revised version of the earlier editions ; it lacks none of the virtues of the earlier texts but, in presenting a fuller account of many areas, gives a greater perspective to the problems.

In his introduction the author clearly defines the limits of his approach ; *nevertheless, the approach is a broad one which aims, successfully, at dealing with the biology of mental defect.* The volume opens with a penetrating historical introduction which covers, among other things, mental measurements, nomenclature, questions of social incompetence, and legislation. The discussion of social incompetence reveals, early in the book, something of the author's attitude : "To the sociologist it may be an advantage to have a social criterion for defining a group, but to the biologist or psychiatrist such a definition is extremely unsatisfactory. Social criteria are not only changeable ; they are relative, not absolute."

The incidence of defectives, classified by age, sex, race and so on, together with the principles of classification, are fully treated in two chapters. Here, as elsewhere, the author's skill as a statistician is apparent. Perhaps the most important conclusions relate to the fertility of severe and mild defectives and the intellectual status of their parents and sibs.

Much of the rest of the text is concerned with a broad treatment of the heredity of mental defect ; a more than adequate introduction on the processes of human heredity is provided. Particular syndromes, expressed at stages varying between the foetal stage and adult life, are fully described ; their hereditary bases are also discussed and these data provide an interesting comparison with later chapters where a variety of environmental effects on mental status is discussed. The author has given a critical summary of the methods of analysis used in human genetics ; this includes consideration of the problems of assortative mating, consanguinity, multi-factorial inheritance and twin analysis. Penrose's treatment of analysis of twins is particularly welcome ; it shows how twins may be exploited in an attempt to disentangle nature and nurture, but it also indicates the limitations and pitfalls which this approach can hold. The important question of the possible decline of intelligence is rightly considered as a problem in population genetics, a problem which must take account of heterosis and the need to maintain genetic flexibility in the population.

This volume is essential reading for anyone interested in human genetics or in mental defect. It combines a thorough scientific approach to the data with a degree of humanity which is uncommon. This latter quality is well illustrated in the quotation : "The ability of a community to make satisfactory provision for its defectives is an index of its own health and progressive development ; the desire for their euthanasia is a sign of involution and decline of human standards." Haldane has provided a characteristically stimulating and provocative preface. Readers will agree with Haldane's view that this book will "... be recognised as a contribution both to thought and to humanism."

J. A. ROPER.

SANDVEN, J., Edit. (1963). *The Role of Educational Research in Social Education*. Oslo : University Press, 1963. Distribution Officer in U.K. : 16, Pall Mall, London, S.W.1. Pp. 302, no price given.

This is a report of the third international congress for the advancement of Educational Research which was held in Oslo in 1961. It contains the papers



presented to the Conference, the findings of study groups, and the general conclusions reached by the participants present. The theme selected for special study was the role of educational research in Social Education, a subject chosen not only for its importance in the field of education, but also because it is one which has been partly neglected.

The papers themselves are grouped into three categories: (1) Social Education from a theoretical viewpoint (six papers and also Introductory Remarks by the President on the theme). (2) Research related to Social Education (thirteen papers). (3) Social Education, as it appears in different geographical areas (five papers). The papers are printed in English, French or German, the three languages of the Conference, with useful summaries at the conclusion of each.

This Conference was, perhaps, unique in that no papers were read during the Conference itself. The contributions had been requested at an early date, were mimeographed, and distributed beforehand to all the participants. This allowed six discussion groups to be arranged during the period of the Conference, and these ran concurrently. They were formed according to the interests of the members and the language most favoured. At these discussions which were based on the contents of the printed articles, the authors could supplement their papers if they wished, and answer questions about them, and of course, other viewpoints could be put forward by any member of the group. The Chairman of the groups reported the main findings at meetings of the plenary sessions and further discussion took place.

The printed volume is of value, therefore, not only because it contains in full the papers themselves, but also for the detailed reports of the discussions of the groups, and the general conclusions reached by the Conference as a whole. The report presents to the reader a picture of the educational thought of the present-day so far as social education is concerned. It shows the trend of opinion of those engaged in this field and the type of research in progress. The findings are not the reflection of any one nation or university, for at the Conference thirty-five nations were represented, and seventy universities. The conclusions reached are to be regarded as efforts to answer the various questions raised, and are not to be considered in any way as final statements.

It is invidious to select any one paper as of outstanding interest or value. Some of the papers are more general than others, particularly those concerned with theoretical problems. The more specific ones are those reporting results of research. One minor criticism is the lack of references in many of the papers. A short list at the end of each paper or at the end of each section or even a brief bibliography would be advantageous to the general reader and to the student of education. This is a volume well worth reading and it can be recommended.

MARY COLLINS.

SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION, XLVIII (1963). *The Scottish Scholastic Survey*, 1953. University of London Press, pp. 216, 25s.

This volume is worthy of the best traditions of the Scottish Council for Research in Education. It completes the Scottish Scholastic Survey which began in 1953, and provides national norms (1953) in Mechanical Arithmetic, Arithmetical Reasoning, English Usage and English Comprehension, for children aged 9 years 11 months to 10 years 10 months. The comparison between the sexes is of interest. In Mechanical Arithmetic, girls and boys are approximately equal except for rural areas, where girls are slightly superior. In Arithmetical Reasoning, boys have a decided lead, but girls are superior in both tests of English. In comparing the four types of area, Cities (1), Large Towns (2), Small Towns (3) and Other Areas (4), it was found that the attainment levels in all four tests followed closely the 3, 1, 2, 4 order established in the mental survey of 1947. In addition, the difference between boys and girls is considerably greater in type 4 than in any of the other three areas; the authors are at a loss to account for this, as it does not occur in all counties classed as Type 4. In particular, it occurs only in one of seven counties of this type, termed 'mainly heath and moor.' The reviewer suggests that this greater sex difference may be



occurring in counties which are more agricultural than the above seven, and that the one apparent exception may be a 'border-line' classification. In such counties, boys might be wanted to work on the farms more often than girls, particularly at busy periods; this might show up in comparative figures of school attendance, i.e., by comparing any sex difference in attendance in the five counties with the difference in attendance (if any) in other regions.

Other interesting findings are that the standard of attainment in the one-teacher schools is as high as that attained by pupils in schools with more than six teachers. In these tests of the short answer type there was no significant difference between the scores of left-handed writers and those of the others, though four out of four differences were all slightly in favour of the right-handed writers.

By no means the least valuable part of the survey is the item by item analysis of the test results and the consequent practical suggestions to the teaching profession about the specific topics or methods in Arithmetic and English which require more time and care or outright change of teaching method.

R. R. DALE.

SMALLWOOD, R. D. (1962). *A Decision Structure for Teaching Machines*. Massachusetts Institute of Technology, Cambridge, Mass., U.S.A.: M.I.T. Press, pp. 122, 30s.

Much of the current literature on teaching machines discards questions of practical evaluation in favour of theoretical gambling. Dr. Smallwood's book is thus the more to be welcomed because of his genuine attempts to parallel working conditions with possible abstract models for developing the automated teaching field. The reader is kept constantly aware of the need for this parallelism, with thorough discussion of the difficulties involved. At no point is there that calculated flight into mathematical operations which so often signifies escape from the more pressing problems of reality and human existence.

The simpler teaching machines in present-day use pay no attention to the individual characteristics and progress of the learner beyond the assessment of such aspects as the errors made and the end-point reached. Some machines offer alternative routes through the material, but the choice of route lies with the learner. Dr. Smallwood describes the theory, construction and practical use of a computer-operated machine which selects the route to be followed by a particular learner after taking account of his own past performance and that of his fellows. A complete system of *machine-decision* is outlined, and an account is given of the underlying probability theory. This latter could be omitted on a first reading. It is in any case beyond all but the mathematically well-informed, and it would indeed be a pity if other interested persons were thus prevented from reading the remaining sections. Here the author describes an actual machine developed from the earlier theoretical basis and used to teach an experimental course. The results show clearly the decisions taken by the machine on the order of presentation of the material, as well as the variations of this order for different students.

This work represents substantial progress in the development of teaching machines, and raises many issues. It demonstrates one way in which automated teaching can move towards the more flexible human approach and yet remain some distance from it—especially since the manner of machine-decision might in some ways be more appropriate to the validation of programmed material than to the prediction of ideal learning sequences. It raises the question, too, of economic *versus* practical considerations, for the cost of such a machine could be out of proportion to its usefulness. Nevertheless, it points the way for important research developments by demonstrating ideas and asking appropriate questions with considerable clarity of vision.

D. M. LEE.



STEVENSON, H. W. (Edit.) (1963). *Child Psychology*. The 62nd Year Book of the National Society for the Study of Education, Part I. Chicago: University of Chicago Press, pp. 550, 48s.

The rate of development of child psychology during the last two decades made it seem unlikely that there would again be a single comprehensive volume to replace Carmichael's *Manual of Child Psychology* which was published in 1946 and reprinted with minor revisions in 1954. This volume cannot claim quite such exhaustive coverage, there is nothing on abnormal or gifted children, for example, only a little on intelligence and no special section on methodology; it does, however, reflect present trends and gives extensive coverage to those areas where most vigorous work is going on. It may thus be regarded as a successor to the *Manual* as this was to the *Handbook* of Murchison.

The general tendency of child psychology to move away from the 'natural history' approach with its somewhat indiscriminate data collection, observational methods, longitudinal and normative studies towards a study of more abstract psychological processes and behaviour constructs, often by means of short-term experimental studies of single variables, has continued at an accelerated pace. The major contribution to this acceleration comes, of course, from the Geneva school, but there are other major influences from learning theory and field theory, for example, and from such outside disciplines as cultural anthropology and sociology.

The changes in emphasis in recent experimental work are carried into the chapter headings of this new book, for half of them were not represented, or only slightly so, in the *Manual*. Of course, a central core of traditional topics, such as learning, perceiving and thinking, remain, but the book starts with two chapters on the biological and sociological correlates of child behaviour. Later, there are four chapters on specific aspects of behaviour; dependence, aggression, achievement and anxiety. The final chapter deals with the theoretical background to modern child psychology and outlines a path by which developmental theory may "undergo transition to a new and more mature level of cognitive organisation."

The writing of the individual chapters is of a uniformly high standard and there is very little overlap among them. The extent and volume of the American research referred to may surprise some workers in this country, but the bibliographies do not contain only American publications. Apart from extensive references to Piaget and his co-workers, notice is also taken of researches in this country, in Russia and other continental countries. It is somewhat chastening to see how small is the British contribution in this field.

This is an important book. It will be too difficult for most students in colleges and departments of education, but it should be in their libraries and in those institutes of education for the use of staff and research workers and advanced students.

L. B. BIRCH.

THOMAS, C. A., DAVIES, I. K., OPENSHAW, D., and BIRD, J. B. (1963). *Programmed Learning in Perspective*. Published for Lamson Technical Products, Ltd., by City Publicity Services, Ltd., London, pp. 182, 28s.

Four members of the Royal Air Force School of Education have written this book "to provide a system of programming for our colleagues in the teaching profession who may wish to write programmes for their own students." They start by surveying briefly the history of programmed learning, the main contemporary types of programme, and their role in teaching situations. They consider their central task of presenting a recipe for programme construction. They illustrate these detailed step-by-step procedures which may be followed and illustrate these considerations by taking the reader through the stages of devising two programmes, one on elementary electricity and one on Pythagoras' Theorem. A final chapter discusses aspects of programming for language training, and there is a selected reading list on programmed learning. The authors would be the last to claim that their recipe is



the final word on programme writing. But they have done well what they set out to do. Their recipe is flexible, yet rigorous. It is presented in lucid, jargon-free language and, so far as the reviewer can judge, embodies many sound principles of learning. In brief, the book provides an outline system by which to construct new programmes and against which to compare existing programmes. Anyone who wants to start writing his own programmes will find this guide useful. It is a tribute to the book's aims and merits that the reviewer would want it to pursue two further questions. Once written, how is the programme used? And how are its effects assessed? This last question is especially important for the whole technology of programmed learning. It is also important for psychology in general. Any system for writing programmes is a prescriptive theory about how people learn and, in some degree, about how they think and communicate with each other. It is a kind of theory which, in comprehensiveness and predictive power, promises more than anything psychology has had before. The loss would be great indeed if we forgot that this kind of theory, no less than any other, stands in need of empirical validation.

I. M. L. HUNTER.

TORRANCE, E. P. (1963). *Education and the Creative Potential*. Minneapolis: University of Minnesota Press. London: Oxford University Press, pp. 167, 36s.

Dr. Torrance is director of educational research at the University of Minnesota: his research programme has the title "The Minnesota Studies of Creative Thinking." Seven general talks on this theme, given between 1960 and 1962, are published here as Part I of the book. Part II presents six precise accounts of experimental studies on specific aspects such as "Peer sanctions against highly creative children" or "Sex-role identification and creativity."

Whom is this book for? The lectures were given to different audiences: administrators, teachers, psychologists, parents. The university or college lecturer might use the first part as an example of the art of the public lecture, especially the skilful use of illustrative case material to support theoretical discussion. The topics of the experimental studies are of the kind normally dealt with at the level of the master's degree, and the reports could serve as models for research students or as source material for other research workers. The variety of approach ensures a comprehensive treatment of the subject, although anyone who gives seven public lectures on one theme in three years, must repeat himself occasionally. For those who are already familiar with the subject from an introductory text, this book will provide good supplementary reading.

The social psychologist will find here a wealth of examples of American social norms and stereotypes: the precocious boy who outsmarts his rigid unimaginative woman teacher; the mother who knows that her child is gifted although he misses his B average ("He just can't harness his brain to such unimportant things... when he's hot on the trail of perpetual motion," page 69); concern with pressures towards conformity; suspicion of conventional academic values. If we find that the image of creative youth is not altogether attractive by European norms, we should perhaps consider, as Torrance asks us to do, whether our educational values are not weighted in favour of conformity and against originality.

JOHN NISBET.

TYLER, L. E. (1963). *Tests and Measurements*. New Jersey: Prentice-Hall, pp. 116, 12s.

This paperback volume is one of the Prentice-Hall Foundations of Modern Psychology Series, the aim of which is to replace the general introductory textbook with a number of short self-contained books, on basic topics. It should be said at once that *Tests and Measurements* attempts to convey no more information than do the relevant sections of some general texts of educational psychology; Professor Tyler's specific aims being to stress simply the concepts needed to read a journal article, to choose a test to serve a particular purpose and to interpret individual test scores.



These aims are attained in three chapters concerned with basic statistical ideas and their application to psychological tests, followed by three chapters on tests of intelligence, aptitudes and achievements and personality, with a final chapter on the application of tests and measurements. There is a short list of selected readings and an adequate subject-index.

In accord with the author's aim of stressing 'consumer' rather than 'producer' knowledge, there is no advice given on such matters as suitable methods of tabulating data nor, for example, is there any attempt to consider the calculation of a mean from grouped data or by assuming an arbitrary mean. Similarly, there is no mention of Spearman's Rank-Order Method of calculating a correlation coefficient. Nevertheless, the essential principles underlying psychological measurement are clearly explained, the treatment of the concepts of validity and reliability being particularly well developed. The section on factor analysis omits any reference to British work, apart from the pioneer contributions of Spearman. A misprint confuses the example of the computation of a *t*-statistic and the term 'stanine', is introduced in a table some twenty pages before the term is explained in the text.

The classification of different types of tests with a consideration of the purposes and limitations of each is straight-forward and acceptable, whilst the detailed description of one or two major tests of each type will doubtless be helpful. The omission of any reference to tests constructed in Britain is one obvious disadvantage for students in this country.

The omissions referred to make it unlikely that this book will be widely used in this country but it is worth consideration because it does appear to be a book that virtually all students could read and understand without tutorial assistance and one which makes a valiant effort to stress the need for "clear thinking and a dash of cautious scepticism" on the part of all who make use of tests and measurements.

E. FISK.

WYNNE, JOHN P. (1963). *Theories of Education*. New York: Harper and Row, pp. xvi+521, 49s.

The author takes twelve of the principal theories of education, with some of their variants, and while giving a substantial account of them, subjects them to a critical appraisal. In each chapter the first two sections are founded on the writing of the proponents of the theory, and the third is a cultural evaluation. When he deals with the older theories he often includes an appraisal of recent scholarly foundations. Other sections of each chapter are devoted to the philosophical aims, and psychological beliefs, beliefs about knowledge and value, educational aims, and practical applications of each of the theories.

Beginning with the hoary but persistent Formal Discipline theory ('The primary aim of education and the good life is the discipline, exercise, or development of the mental faculties'), we are led through the Natural-Perfection theory of the newly-discovered Rousseau, past Herbart and the Appreciation theory and James' Habit Tendency theory to Thorndike's 'Improving and satisfying human wants.' Dewey, the 'philosopher of growth' defines education as the 'improvement of the quality of experience,' while insisting on the importance of democracy. There is a logical continuation of this in Kilpatrick's democratic character theory, followed by Bode's thesis of 'democratic growth.' Bayles, a disciple of Bode, rejects any belief that democracy requires either "letting students . . . do largely as they please or . . . letting them decide what shall be the rules of behaviour." Like Bode, however, he "would enlist their participation in school management insofar as they are capable, and permit them to make mistakes." Also of interest to psychologists is Bayles' conception of educational progress, which "should be conceived in terms of some such ends as improvement in independent thinking, expansion of outlook, and integration of values. Consequently, the estimation of progress would have to be shifted from emphasis on the acquisition of isolated facts and skills to improvement in reflective thinking and to the adequacy and harmony of outlook."

After discussing the ideas of Professor Horne on the educational implications of absolute idealism the author gives an account of the Supernatural Development theory founded in the Roman Catholic faith. Later chapters deal with theories founded on 'The Great Classics, Social Self-Realisation, Self-Perception,' and the 'Basic Education' of the new conservatives.

The material is presented logically and clearly, and the author makes a thorough job of his task, so that his book will be of value to advanced students of education. Presented in this way, however, it is by no means the right fare for students training to be teachers, unless taken in very small doses. As one reads, one also begins to wonder whether, among so many philosophies of education that are faulty, there will ever emerge a philosopher who will evolve one on which we can rely. Meanwhile, *cui bono*?

R. R. DALE.

## NEW JOURNAL

### THE TOHOKU JOURNAL OF EDUCATIONAL PSYCHOLOGY

*Editor* : KINJU MATSUMOTO.

We welcome this Japanese edition to the world's journals of educational psychology, which began in May and is to appear twice a year from now on, in May and October. The journal will act mainly as an outlet for Japanese workers, but will also publish papers from others. The first number contains six local authors and one Russian. It is published wholly in English, but a note with the journal states that it will be published in English, German or French. It is not quite clear whether this implies that some future numbers may be in any one or a mixture these languages, or whether journals for this country will always be in English.

The coverage would seem to be a little wider than the *British Journal*, as it includes abnormal, clinical and industrial psychology. The first number contains a survey of Soviet psychology, an account of experiments to evaluate some of Piaget's findings, a study of lip-reading, and a number of case histories of aphasic children.

The journal may be obtained from The Tohoku Association of Educational Psychology, c/o Institute of Educational Psychology, Faculty of Education, Tohoku University, Sendai, Japan.



## PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

- L. S. PENROSE (1963). *The Biology of Mental Defect*. Sidgwick and Jackson, Ltd., pp. xxiv+374, 42s.
- P. L. BROADHURST (1963). *The Science of Animal Behaviour*. Penguin Books, pp. 142, 3s. 6d.
- ERICH FROMM (1963). *The Dogma of Christ*. Routledge and Kegan Paul, Ltd., pp. viii+151, 18s.
- MARGARET DONALDSON (1963). *A Study of Children's Thinking*. Tavistock Publications, pp. viii+263, 35s.
- PRYNCE HOPKINS (1963). *Orientation, Socialization and Individuation*. Asia Publishing House, pp. 100, 27s. 6d.
- LOREN EISELEY (1963). *The Mind as Nature*. Harper and Row, Ltd., pp. 60, 20s.
- HAROLD W. STEVENSON (Edit.) (1963). *Child Psychology*. Chicago: The National Society for the Study of Education, pp. x+550+vi, 48s.
- B. M. FOSS (Edit.) (1963). *Determinants of Infant Behaviour II*. Methuen and Co., Ltd., pp. xii+248, 45s.
- E. M. RENWICK (1963). *Children Learning Mathematics*. Arthur H. Stockwell, Ltd., pp. 94, 9s. 6d.
- CARL H. DELACATO (1963). *The Diagnosis and Treatment of Speech and Reading Problems*. Illinois: Charles C. Thomas, pp. x+188, \$6.75.
- RICHARD H. BLOOMER (1963). *Reading Comprehension for Scientists*. Illinois: Charles C. Thomas, pp. xiii+213, \$8.75.
- P. J. LAWRENCE (Edit.) (1963). *Mental Health and the Community*. New Zealand: Canterbury Mental Health Council, pp. 590, 35s.
- ANDRE BERGE (1963). *The Sexual Education of Children*. Sheed and Ward, pp. x+150, 9s.
- GILBERT E. TEAL (Edit.) (1963). *Programmed Instruction in Industry and Education*. Stamford, Conn., U.S.A.: Public Service Research, Inc., pp. ix+316, \$7.50.
- HERBERT C. QUAY (Edit.) (1963). *Research in Psychopathology*. New Jersey: D. Van Nostrand Co., Inc., pp. viii+216, 15s. 6d.
- CHARLES H. SOUTHWICK (Edit.) (1963). *Primate Social Behavior*. New Jersey: D. Van Nostrand Co., Inc., pp. vi+191, 15s. 6d.
- ROBERT F. GROSE and ROBERT C. BIRNEY (Edits.) (1963). *Transfer of Learning*. New Jersey: D. Van Nostrand Co., Inc., pp. vi+194, 14s.
- ROBERT A. BAKER (Edit.) (1963). *Psychology in the Wry*. New Jersey: D. Van Nostrand Co., Inc., pp. xii+170, 14s.
- JOHN R. BRAUN (Edit.) (1963). *Contemporary Research in Learning*. New Jersey: D. Van Nostrand Co., Inc., pp. vi+229, 15s. 6d.
- EUGENIA HANFMANN, RICHARD M. JONES, ELLIOT BAKER and LEO KOVAR (1963). *Psychological Counseling in a Small College*. Massachusetts: Schenkman Publishing Co., Inc., pp. xi+131, \$3.95.
- THE SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1963). *The Scottish Scholastic Survey, 1953*. University of London Press, Ltd., pp. 216, 25s.
- H. A. DAVID (1963). *The Method of Paired Comparisons*. Charles Griffin and Co., Ltd., pp. 124, 28s.
- THEODORA ALCOCK (1963). *The Rorschach in Practice*. Tavistock Publications, pp. xii+252, 63s.
- JEAN THOMAS and JOSEPH MAJAUULT (1963). *Primary and Secondary Education: Modern Trends and Common Problems*. Council for Cultural Co-operation of the Council of Europe, pp. 140, 10s.
- HERTHA ORGLER (1963). *Alfred Adler: The Man and his Work*. Sidgwick and Jackson, Ltd., pp. xix+266, 30s.
- ALBERT BANDURA and RICHARD H. WALTERS (1963). *Social Learning and Personality Development*. New York: Holt, Rinehart and Winston, Inc., pp. ix+329, \$7.00.

- DONALD A. SCHON (1963). *Displacement of Concepts*. Tavistock Publications, xiii+208, 30s.
- STEN HEGELER (1963). *Choosing Toys for Children*. Tavistock Publications, xx+127, 21s. 6d.
- N. D. SUNDBERG, and L. E. TYLER (1963). *Clinical Psychology*. Methuen and Co., Ltd., pp+xviii+564, 50s.
- JULIUS E. HEUSCHER (1963). *A Psychiatric Study of Fairy Tales*. Springfield, Illinois; Charles C. Thomas, pp. x+224.
- E. PAUL TORRANCE (1963). *Education and the Creative Potential*. Oxford University Press, pp.vii+167, 36s.
- A. M. DUSHKIN and C. FRANKENSTEIN (Edits.) (1963). *Scripta Hierosolymitana*, Vol. XIII, Studies in Education. Jerusalem: Magnes Press, Hebrew University, pp. 264, 43s.
- HENRY JOHNSON (1963). *Sound Way to Correct Spelling*, Stages 1, 2 and 3. Longmans, Green and Co., Ltd., Stage 1, 3s. 3d., Stage 2, 3s. 6d., Stage 3, 3s. 9d.
- M. M. LEWIS (1963). *Language Thought and Personality*. George G. Harrap and Co., Ltd., pp. 256, 18s.



## NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be preceded by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need *not* be included, but successive sections of the article should receive numbered headings, e.g. :

- I. INTRODUCTION.
- II. OUTLINE OF PREVIOUS RESEARCH.
- III. PLAN OF THE INVESTIGATION.
  - (i) *Subjects and materials.*
  - (ii) *Experimental Procedure.*
- IV. STATEMENT OF RESULTS.
- V. DISCUSSION AND CONCLUSIONS.
- VI. REFERENCES.

4.—*Typing.* All papers must be typed (double or 1½-spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs \*, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.<sup>1</sup> When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—  
for any one author—by order of date. The following standard form of references should be observed :

DEARBORN, W., F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development*. Cambridge, Mass. : Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)  
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71.

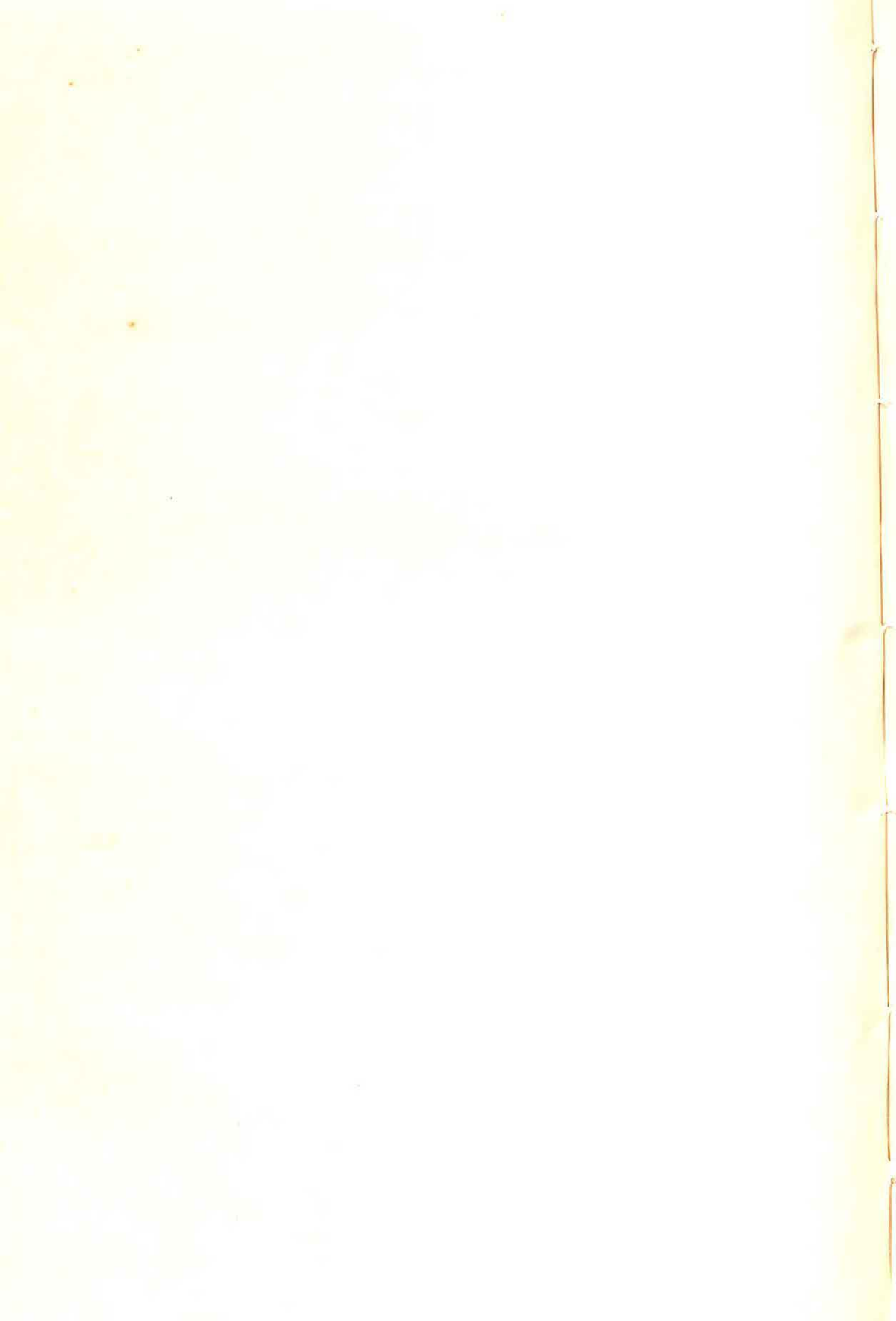
(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full).

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table and a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in the text. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9.—*Reprints.* Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty may be had at cost price if the order is given when the proof is returned.





ADAPTATION TO LEVEL OF DIFFICULTY IN JUDGING  
THE FAMILIARITY OF WORDS

BY A. W. HEIM, B. RAMSAY AND K. P. WATTS

*(Psychological Laboratory, Cambridge)*

**SUMMARY.** The effect of context-difficulty on judging the degree of familiarity of obscure words was investigated. Three experiments are described in which the subjects were, respectively, students from two different universities and naval ratings. All three groups displayed unwitting adaptation to the level of difficulty of the context in which the words were presented. Thus subjects, who were required to judge the familiarity of forty selected obscure words presented together with forty relatively easy words tended to judge the obscure words as unfamiliar; whereas those subjects (equated on vocabulary strength) who were presented with the selected words, together with forty extremely difficult words, tended to judge the selected words as familiar—the set of selected obscure words in the two cases being, in fact, identical. (Different sets of words were used, of course, for the students and for the naval ratings.)

In addition, certain general differences emerged between the attitude of university students to little-known words, and the attitude of naval ratings. The former were most influenced by context in their judgments of extreme *unfamiliarity* whilst for the latter, the context exerted a stronger effect on their judgments of complete *familiarity*.

## I.—INTRODUCTION.

It has been shown that adaptation to level of difficulty occurs in intelligence testing (Heim, 1955), i.e., a group's standard of performance on an intelligence test is determined partly by the adaptation level of the group members. Previously equated subjects have been found to solve 'difficult' problems significantly more often when these problems were presented in a context of equally, or more, difficult problems than when the background provided (for identical problems) was relatively easy. The extent of the adaptation increased with the proportion of difficult items in the test; the process was found to occur not merely within a test but also to carry over from one test to another taken immediately afterwards. Such adaptation occurred only within limits prescribed by the relationship of test-level to group-level: it did not appear if the crucial test items were so hard for the subjects that they were scarcely able to score above chance.

Unwitting adaptation of this kind has been observed and investigated by many experimental psychologists, outside the psychometric field. Helson, for example, has done a great deal of work (1938, 1947, 1948, 1949), using variation in stimulus intensity, as opposed to variation in the difficulty of a task. He has put forward a theory to cover psychophysical data (such as weight-lifting and colour vision) and responses to repetitive tasks, and he suggests that the theory may be usefully applied also to social, moral and aesthetic behaviour, both group and individual.

Mowbray (1953) carried out experiments on the simultaneous presentation of visual and auditory stimuli. He used prose passages of three different levels of difficulty and demonstrated, among other things, that with simultaneous presentation, significantly greater deterioration occurred with the easier than with the harder passages, i.e., subjects tended to retain the former less well

than the latter. This finding is consonant with that of Szafran and Welford (1950) that: "If two tasks of unequal difficulty are presented, achievement at the two tasks considered together, after allowance has been made for practice effects, will tend to be greater when the harder is presented first, than when the easier is presented first." Gibbs (1951) reached a similar conclusion. Mace's work on 'implicit standards' and 'indirect incentives' (1931, 1935) again suggests that altering the standards set up for a subject—either by changing the instructions or by giving knowledge of results—effects a predictable alteration in the subject's absolute performance.

Heim has elsewhere discussed at some length these experiments, and others on 'level of aspiration' (Heim, 1957). The 'level of aspiration' investigations, however, unlike the others quoted, concern the solicited expression of an opinion on the part of a subject as to his performance. In this respect, these latter experiments somewhat resemble those to be described in this paper. The aim of the work described below was to ascertain whether subjects' *judgments of the degree of familiarity* of given words vary with the difficulty of the context in which these words are presented.

## II.—THE TESTS.

The tests used were drawn from the Self-judging Vocabulary Scale (Heim and Watts, 1961). This scale consists of five sets of forty words named, respectively, A, E, I, O, and U. The level of difficulty of the words has been ascertained experimentally. A is the easiest, E is slightly harder, I is medium in difficulty, O is harder and U consists of extremely difficult and obscure words. Each complete test comprises eighty words, made up of two adjoining sets. For example, 12 and 13-year-old school children would take AE; adults who left school at 15 or earlier would take EI; sixth-formers and most students would take IO; university graduates would take OU.

The test is conducted in two parts. The subjects are first presented with their list of eighty words, e.g., AE. These words are arranged alphabetically in five columns of eight words (set A) in the top half of the page, and another five columns of eight words (set E) in the lower half of the page. For the benefit of the scorer, set A is separated from set E by a small space (imperceptible to the subject).

The instructions are as follows: "Please write an *a*, *b* or *c* on the left of every word. Write *a* if you know the word and could explain its meaning to someone unfamiliar with it; write *b*, if you are doubtful; and *c*, if you have never seen the word before and have absolutely no idea of its meaning. Begin at the top of the column on the left and work right down to the bottom of each column in turn. When you have finished, look through to make sure you have not missed out any and then put up your hand."

As each subject finishes, his list is removed and he is handed Part II of the test. This consists of a booklet with the same words in the same order (A pages and E pages alternating), but after each word there are printed six phrases. The subject is told that in every case one of the six phrases corresponds to the meaning of the crucial word, the other five being incorrect. He is asked to put a ring round the number of the phrase which, in his view, most closely denotes the meaning; to take every item in turn, guessing where he does not know; and to give a seventh answer in his own words when he does know the crucial word but dislikes all the alternatives offered. The position of the correct solution is randomised. No time limit is imposed.



Previous work on the Self-judging Vocabulary Scale has shown : (i) That it is exceedingly rare for a subject to get right 100 per cent. of the words he marked as *a*. (ii) That the majority of subjects get right more than the one-in-six of their *c* words, which would be expected by chance. (iii) That the absolute number of *a*'s and *c*'s given by a subject varies concomitantly with his score, i.e., those with high scores on the SJ Vocabulary test tend to give a high number of *a*'s (and those with poor scores tend to give many *c*'s and few *a*'s). But (iv) the number of *b*'s given bears no obvious relation to the Vocabulary score of the subject.

### III.—AIMS.

The main question posed in this enquiry was whether the *abc*'ing of a list of given words would be influenced by the context in which these words were presented. If, for example, the O list of words (hard) were chosen for experimentation, then these O words could be presented to one group of subjects along with I words (of medium difficulty) and to another, equated, group of subjects along with U (very hard) words. Our hypothesis was that the O words in the IO list would have fewer *a*'s and more *c*'s than would the (identical) O words in the OU list. That is, the IO subjects would 'adapt downwards,' judging the O words as less familiar, in view of the relative ease of the accompanying I words. Whereas the OU subjects would 'adapt upwards,' judging the O words as more familiar, in view of the relative obscurity of the accompanying U words.

It was important that the two groups of subjects be equated for vocabulary strength, in view of the tendency already observed for people with higher scores on the Vocabulary test to give a higher proportion of *a*'s than do people with lower Vocabulary scores. Experiment 1 was, therefore, conducted as follows.

### IV.—EXPERIMENT 1.

The subjects were forty Cambridge University undergraduates. All were male and were reading either Natural Sciences or Mechanical Sciences. These subjects were from the beginning divided, arbitrarily, into two subgroups of twenty, an IO subgroup and a OU subgroup. Every subject was handed a sheet on which were listed eighty words, but subgroup IO had forty words of medium difficulty (designated as I words) and forty hard words (designated as O words); subgroup OU had forty hard words (the same O words) and forty very hard words (designated as U words). For both subgroups the words were arranged so that, as the subject worked through his *abc* sheet, he alternated (eight words by eight words) between the relatively easy and the relatively difficult words, as described in Section II.

The subjects were tested in small groups, IO members together and OU members together. They received the usual *abc* instructions (as quoted above). When they had finished, the *abc* sheet was removed and the subjects were given a question book in which only the O words were listed, i.e., the forty 'hard' words which were common to both of the *abc* sheets. Each of the O words in the question book was followed by six phrases of which one expressed the meaning of the given word. The procedure at this stage was the usual one for the SJ Vocabulary test.

Thus, all the subjects of subgroup IO and subgroup OU were tested on their knowledge of the O words. The maximum possible score was 40 marks. On the basis of their scores, it was possible to select fifteen pairs of subjects,

each member of the pair scoring (within two marks) the same as his opposite number in the other subgroup. The total O Vocabulary score of the fifteen IO subjects equals that of the fifteen OU subjects (288). Thus, the two subgroups used were matched for actual word knowledge.

TABLE 1  
CAMBRIDGE UNIVERSITY STUDENTS.  
Scores on Set O.

No. of Pair	IO Subjects	OU Subjects
1	25	25
2	24	24
3	23	22
4	23	22
5	22	21
6	21	21
7	21	21
8	19	18
9	19	18
10	19	18
11	17	17
12	17	17
13	14	16
14	12	14
15	12	14
Total .....	288	288
Mean .....	19.2	19.2
s.d. ....	4.0	3.3

Table 1 shows the matching arrangement chosen. It may be seen that as a matter of convenience, the numbering of the pairs goes simply from top scorer (1) to lowest scorer (15). It may be seen, too, that several pairs could, in fact, have been differently matched. Subject 3 of subgroup IO, for instance, could have been paired with subject 4 of subgroup OU. When analysing the *a*, *b* and *c* results, these possibilities were examined, however, and it was found that such rearrangement effected virtually no change in the results.

The following *a*, *b* and *c* scores were obtained from the fifteen pairs.

Three points of interest may be seen in Table 2.

(i) The first concerns the number of *c*'s given by the two subgroups. Subgroup IO subjects gave a total of 271 *c*'s, yielding a mean of 18.1, whereas the corresponding figures for the OU subjects were, respectively, 207 and 13.8. As our experiment involved paired observations, we calculated the standard error of the mean difference (Edwards, 1946) and hence the value of *t*. This gave  $t=3.41$  which is significant at the 0.0025 level, using a one-tailed test. Thus, on the *c*'s, the result is in the predicted direction, at a highly significant level of confidence. The IO subjects judged significantly more of the O words on the *abc* sheet as totally unfamiliar than did the (equated) OU subjects.

(ii) The second point concerns the number of *a*'s given by the two subgroups. Here again, the difference between the IO and OU subjects was in the predicted direction, namely that owing to upward adaptation, the OU subjects



TABLE 2  
CAMBRIDGE UNIVERSITY STUDENTS.  
Results on Set O.

No. of Pair	IO Subjects			OU Subjects		
	Total <i>a</i>	Total <i>b</i>	Total <i>c</i>	Total <i>a</i>	Total <i>b</i>	Total <i>c</i>
1	16	15	9	23	12	5
2	16	9	15	21	10	9
3	17	9	14	21	5	14
4	22	5	13	10	13	17
5	11	9	20	25	6	9
6	19	0	21	23	8	9
7	10	15	15	15	13	12
8	15	5	20	11	10	19
9	16	7	17	15	11	14
10	14	7	19	12	13	15
11	14	7	19	15	17	8
12	12	8	20	12	14	14
13	9	7	24	15	7	18
14	8	9	23	7	6	27
15	8	10	22	11	12	17
Total .....	207	122	271	236	157	207
mean .....	13.8	8.1	18.1	15.7	10.5	13.8

would mark their O words as *a* more frequently than would the IO subjects. The respective figures here are as follows: OU total *a*'s=236, mean 15.7; IO total *a*'s=207, mean 13.8. This difference, however, is far from being statistically significant ( $p=0.12$ ). The lack of significance may be attributed largely to subject No. 4 of the OU subgroup. If pair 4 be omitted,  $t=2.473$ ,  $p=0.02$  on a one-tailed test.

(iii) On the 'doubtful' category, the words marked by subjects as *b*, no prediction had been made. In fact, it may be seen that the difference here between the IO and OU subgroups goes in the same direction as it did for the *a*'s: i.e., the OU subjects gave more *b*'s than did the IO subjects.  $t$  here=1.915, yielding a value of  $p$  between 0.1 and 0.05 (two-tailed test).

The wide range and variability of the *b* judgments is interesting. It may be seen, for instance, with Nos. 6 and 7 in subgroup IO, that although they gained an identical O score (of 21), one of them gave no *b*'s at all, whilst the other gave as many as fifteen *b*'s (see Tables 1 and 2).

The results were sufficiently promising to suggest that further work on these lines might be worthwhile. We decided, therefore, to examine with the same aims in mind, two further groups of subjects whom we had already tested. These enquiries will be referred to as Experiment 2 and Experiment 3.

#### V.—EXPERIMENT 2.

The subjects in Experiment 2 were students at the University of Keele, men and women. They were tested in the first week of their first year at Keele. These subjects were divided arbitrarily into two subgroups and the SJ Vocabulary tests used, as in Experiment 1, were IO ( $n=84$ ) and OU ( $n=93$ ). The O mean score for subgroup IO=15.14 ( $\sigma=5.3$ ) and the O mean score for subgroup OU=15.89 ( $\sigma=5.2$ ). These differences are not statistically significant.

TABLE 3  
KEELE UNIVERSITY STUDENTS.

Results on Set O.

IO Subjects			OU Subjects		
mean <i>a</i>	mean <i>b</i>	mean <i>c</i>	mean <i>a</i>	mean <i>b</i>	mean <i>c</i>
9.5	9.5	20.9	11.2	10.0	18.8

Table 3 gives the mean *a*, *b* and *c* results on Set O, the set which was taken by both subgroups. It may be seen that the *a* and *c* differences are in the predicted direction, confirming that the IO subjects adapted 'downwards' and the OU subjects 'upwards' in accordance with the relative ease of the I set and the difficulty of the U set.

Since in this experiment the matched-pairs technique was not appropriate, it was decided to carry out an analysis of variance on these data. This would allow for the variance between scores (since the number of *a*'s and *c*'s are known to vary with the number of SJ words correctly answered) and would provide an estimate of the variance between subgroups owing to adaptation.

TABLE 4  
KEELE UNIVERSITY STUDENTS TAKING IO AND OU TESTS.

Analysis of Variance on Set O.

	d.f.	Mean Sq.	V.R.
Number of <i>a</i> 's given :			
Between scores . . . . .	26	133.307	7.361 $p < 0.001$
Between subgroups . . . . .	1	116.463	6.431 $p < 0.05$
Remainder . . . . .	149	18.108	
	176		
Number of <i>b</i> 's given :			
Between scores . . . . .	26	12.784	0.981 ns
Between subgroups . . . . .	1	20.279	1.556 ns
Remainder . . . . .	149	13.034	
	176		
Number of <i>c</i> 's given :			
Between scores . . . . .	26	108.972	4.439 $p < 0.001$
Between subgroups . . . . .	1	207.119	8.438 $p < 0.01$
Remainder . . . . .	149	24.548	
	176		



The analysis of variance shows the usual relationship between the score gained (on set O) and the number of *a*'s and *c*'s given. As regards the adaptation level, this group yields very satisfactory results as they are significant in the predicted direction for both the *a*'s and the *c*'s. Significantly more *a*'s are given in set O by the OU subgroup than by the IO subgroup; and significantly more *c*'s are given by the IO than the OU subgroup. It may be seen, however, that the *c* effect reaches a higher level of significance than does the *a* effect. Neither adaptation nor score had a significant effect on the *b* judgments.

#### VI.—EXPERIMENT 3.

The subjects in Experiment 3 were naval ratings in their late teens and early twenties. The two SJ tests given in this case were AE ( $n=47$ ) and EI ( $n=105$ ). It may be recalled that set A is easy; E is slightly harder; and I is a set of medium difficulty. The numbers in the two subgroups are not as well balanced as they were in Experiments 1 and 2, since the sailors had been tested at an earlier date, the main aim, at that time, being to establish norms. Their mean scores on E, however, were sufficiently similar to allow us to use these two subgroups for the present purpose. The mean E score for subgroup AE=24.77 ( $\sigma=7.6$ ) and the mean E score for subgroup EI=24.03 ( $\sigma=6.2$ ). The difference between these two means is not statistically significant.

TABLE 5  
NAVAL RATINGS.  
Results on Set E.

AE Subjects			EI Subjects		
mean <i>a</i>	mean <i>b</i>	mean <i>c</i>	mean <i>a</i>	mean <i>b</i>	mean <i>c</i>
16.4	11.5	12.2	18.9	9.9	11.3

Table 5 gives the mean *a*, *b* and *c* results on set E, the set which was taken both by the AE and the EI subgroup. It may be seen that the *a* and the *c* differences are in the predicted direction, suggesting that the AE subjects adapted 'downwards' and the EI subjects 'upwards', in accordance with the relative ease of the A set and the difficulty of the I set, for these subjects. Again, an analysis of variance was carried out, the results of which are given in Table 6.

The analysis of variance shows, as is to be expected, that the difference due to variations in score is highly significant. The higher the score of a subject, the greater the number of *a*'s and the smaller the number of *c*'s given. For these subjects, the variation in the numbers of *b*'s also was found to be related to score on set E. With these score-effects looked after by the analysis of variance, we may now turn our attention to the variance due to differences between the two subgroups. It will be seen that the variance ratios are significant both for the *a*'s and the *b*'s. More *a*'s are given in Set E by the EI subgroup than by the

TABLE 6

NAVAL RATINGS TAKING EI AND IO TESTS.

Analysis of Variance on Set E.

	d.f.	Mean Sq.	V.R.
Number of <i>a</i> 's given :			
Between scores .....	28	160.596	5.751 $p < 0.001$
Between subgroups .....	1	202.180	7.240 $p < 0.01$
Remainder .....	122	27.926	
	151		
Number of <i>b</i> 's given :			
Between scores .....	28	35.262	1.993 $p < 0.01$
Between subgroups .....	1	83.264	4.706 $p < 0.05$
Remainder .....	122	17.693	
	151		
Number of <i>c</i> 's given :			
Between scores .....	28	105.174	3.256 $p < 0.001$
Between subgroups .....	1	25.951	0.803 ns
Remainder .....	122	32.301	
	151		

AE subgroup as was predicted ( $p < 0.01$ ). But more *b*'s were given in set E by the AE subgroup than by the EI subgroup ( $p < 0.05$ ), no prediction having, in fact, been made for the *b*'s. For the *c*'s the variance ratio for subgroups was not significant although, as predicted, the AE subgroup gave more *c*'s on Set E than the did EI subgroup. The lack of significance may be seen from Figure 1 to be due to a small group of very low scorers in subgroup EI.

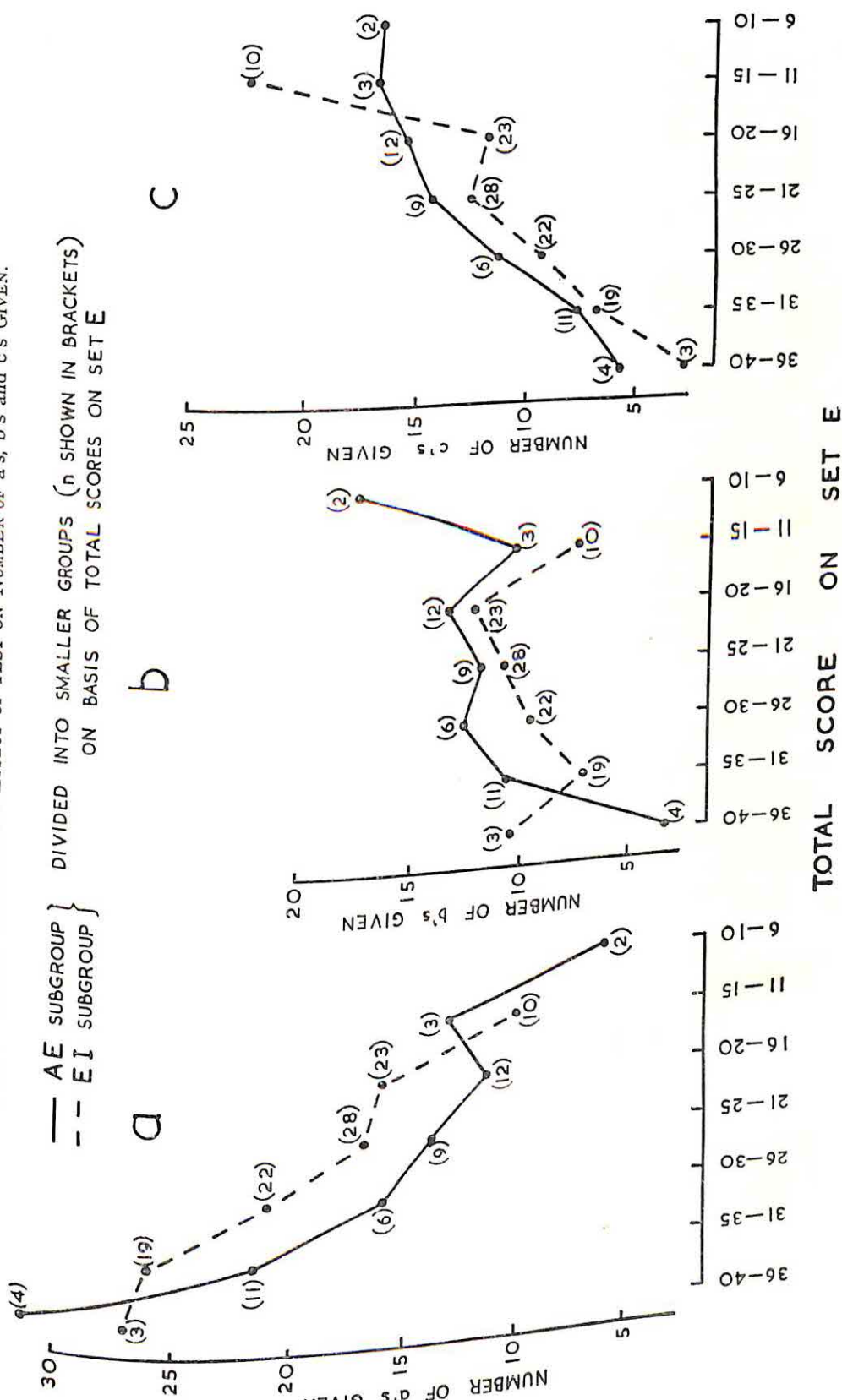
Figure 1 shows the number of *a*'s, *b*'s and *c*'s plotted against the score gained on set E (the maximum possible score being 40). The continuous line represents the AE subjects and the dotted line represents the EI subjects. Thus, the distance between the two graphs, in each case, shows the effect of adaptation to level of difficulty—the greater the interval the greater the degree of adaptation. The most interesting of these pairs of graphs is that for the *c*'s. This shows clearly that the lowest-scoring EI subjects are 'out of place': these ten subjects gave a mean of over 20 *c*'s and gained a mean E score of 15 marks or less. For these few subjects test EI was probably an inappropriate test, in the sense that probably part E and certainly part I, were too difficult for them. (They gained a mean score of 7.5 on I, which is around chance.)

Figure 1 also shows clearly the effect of E score on the number of *a*'s, *b*'s and *c*'s given. The graphs go down for the *a*'s, up for the *c*'s and are, jointly, near-horizontal for the *b*'s.

It is interesting that whilst the adaptation effect was found both with the university groups and the naval ratings group, it was stronger for the former with the *c* judgments and for the latter with the *a* judgments.



FIG. 1 EFFECT OF SCORE AND EFFECT OF TEST ON NUMBER OF a's, b's and c's GIVEN.



## VII.—DISCUSSION.

The main aim of this work was to ascertain whether the adaptation to level of difficulty which has been demonstrated in many psychological fields, takes place with regard to *judgments of the degree of familiarity* of unusual words. The hypothesis was that subjects presented with obscure words in a difficult context would tend to judge these words as more familiar than would subjects (equated for vocabulary strength) presented with the same obscure words in an easier context. This hypothesis has been confirmed for three different groups, using two different tests, both drawn from the Self-judging Vocabulary Scale.

All three experiments yielded positive results but the emphasis is interestingly different, especially as between students (Experiments 1 and 2) and naval ratings (Experiment 3). For both the Cambridge and the Keele students, the *c* adaptation effect reached significance level. With Keele, the *a* effect was also significant but at a much lower level than the *c* effect. Thus, for the university subjects, it was primarily the judging of extreme *unfamiliarity* ("never seen the word before and have no idea what it means") that was most influenced by the context in which the obscure words appeared.

With the sailors, however, it was only the *a*'s whose adaptation results reached significance level: with them, it was the judging of complete *familiarity* ("know the word and could explain its meaning") that was most influenced by the context in which the obscure words appeared. This difference between university students and naval ratings might have derived from differences in group/test-level relationship or it might have been due to the fact that the sailors had a larger standard deviation than either of the student groups. We therefore examined the mean scores gained by the three groups on set O (Cambridge and Keele students) and set E (naval ratings). The respective means were as follows: 19.2 ( $\sigma=3.7$ ), 15.5 ( $\sigma=5.3$ ) and 24.3 ( $\sigma=6.6$ ).

Thus, the relevant vocabulary set (E) was easier for the naval ratings as a group than was O for both of the university groups, but the vocabulary set (O) was easier for the Cambridge students than it was for the Keele students. Yet, the Keele students showed the predicted adaptation effect significantly for both their *a*'s and their *c*'s. This suggests that the difference in emphasis displayed between students and sailors is not due to the relative ease of set E for most of the sailors. (It is, perhaps, worth recalling that for a small minority of the sailors, set E was, in fact, too difficult, see Figure 1, and that earlier work has, understandably, demonstrated lack of adaptation in cases where subjects are scarcely able to score above chance).

The most likely cause for the failure of Experiment 1 to show a significant effect on the *a* judgments, is the paucity of the numbers in this experiment. It may be remembered that the omission of one of the pairs of Cambridge students would have rendered the *a* effect significant.

There is a further point which is linked with this question of greater *a* or *c* adaptation: this concerns the relation of adaptation to *b* judgments. In fact, we had expected little or none. Some effect was found, however, in all three experiments. For the two university groups, the *b* adaptation went in the same direction, namely 'upwards,' as with the *a* judgments: the students tended to give more *b* judgments to O words, when these words were presented in a difficult context (OU) than when presented in an easier context (IO). The effect did not reach significance level.

The *b* effect with the naval ratings did attain significance level and it went in the opposite direction, namely 'downwards,' as with their *c* judgments. These subjects gave more *b* judgments to E words, when these words were



presented in an easier context (AE) than when presented in a harder context (EI). It looks, therefore, as though the 'doubtful' category may be implicitly regarded as closer to the 'unfamiliar' for sailors and as closer to the 'familiar' for students. It is as though students are clearly aware of which words they do *not* know and draw a rather fine line between the wholly familiar and the faintly familiar, whereas less educated subjects are clearly aware of which words they *do* know and draw their fine line between the wholly unfamiliar and the faintly familiar.

The peculiarly individual nature of *b* judgments must be stressed, however. Unlike the *a* and *c* judgments, the *b*'s do not obviously correlate with total vocabulary score (see Figure 1) and they do yield a very wide range for any one score.

The three experiments taken together indicate the functioning of unconscious adaptation to level of difficulty in judging the degree of familiarity of words presented in contexts of varying difficulty. As with the work on adaptation to level of difficulty in intelligence testing, this suggests that the treatment of test scores as absolute may often be misleading and that it is advisable to extend, as fully as possible, subjects in all psychometric situations.

The testing in Experiment 1 was carried out by B. Ramsay as part of her research project for the Part II Psychology Tripos. The other two authors wish to express their thanks to the Medical Research Council for financial support; to Mr. A. Iliffe for his kindness in arranging for the testing of students in Experiment 2; and to all the subjects for their co-operation.

#### VIII.—REFERENCES.

- EDWARDS, A. L. (1946). *Statistical Analysis for Students in Psychology and Education*. New York: Rinehart.
- GIBBS, C. B. (1951). Transfer of training and skill assumptions in tracking tasks. *Quart. J. Exp. Psychol.*, 3, 99-110.
- HEIM, A. W. (1955). Adaptation to level of difficulty in intelligence testing. *Brit. J. Psychol.*, 46, 211-224.
- HEIM, A. W. (1957). Psychological adaptation as a response to variations in difficulty and intensity. *J. Gen. Psychol.*, 56, 193-211.
- HEIM, A. W., and WATTS, K. P. (1961). A preliminary study of the self-judging vocabulary scale. *Brit. J. Psychol.*, 52, 175-186.
- HELSON, H. (1938). Fundamental problems in color vision: I.—The principle governing changes in hue, saturation and lightness of non-selective samples in chromatic illumination. *J. Exp. Psychol.*, 23, 439-476.
- HELSON, H. (1947). Adaptation level as frame of reference for prediction of psychophysical data. *Amer. J. Psychol.*, 60, 1-29.
- HELSON, H. (1948). Adaptation level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 55, 297-313.
- HELSON, H., and JEFFERS, V. B. (1949). Fundamental principles in color vision: II.—Hue, lightness, and saturation of selective samples in chromatic illumination. *J. Exp. Psychol.*, 26, 1-27.
- MACE, C. A. (1931). The influence of indirect incentives upon the accuracy of skilled movements. *Brit. J. Psychol.*, 22, 101-114.
- MACE, C. A. (1935). *Incentives: Some Experimental Studies*. Rep. Industr. Hlth. Res. Bd., Lond., No. 72.
- MOWBRAY, G. H. (1953). Simultaneous vision and audition: The comprehension of prose passages with varying levels of difficulty. *J. Exp. Psychol.*, 46, 365-372.
- SZAFRAN, J., and WELFORD, A. T. (1950). On the relation between transfer and difficulty of initial task. *Quart. J. Exp. Psychol.*, 2, 88-94.

(Manuscript received, September 30th, 1963.)

# A LIE SCALE FOR THE JUNIOR MAUDSLEY PERSONALITY INVENTORY

By H. B. GIBSON

*(Institute of Criminology, University of Cambridge)*

**SUMMARY.** An eighteen item lie scale (L) was incorporated in the Junior Maudsley Personality Inventory (J.M.P.I.) and the combined instrument was administered to 161 English school children. An item analysis of the L scale was carried out separately for boys and for girls. The item-scale correlations for all of the items were satisfactory and no items had significant sex differences. The split-half reliability of the L scale was .65.

The group mean L score of the girls was significantly higher than that of the boys, but there were no significant sex differences in means and S.Ds. for the Neuroticism (N) or Extraversion (E) scales, nor were these data significantly different from the published norms for the test.

A significant negative correlation was found between L and N, but no significant correlation between L and E. These findings are similar to those which have been obtained with the corresponding adult scale, the Maudsley Personality Inventory (M.P.I.), as far as inter-scale relationships are concerned but with the children's Inventory the distributions of L scores are roughly normal, and this has implications for the very low scorers as well as those who score very high. The validity of the concept of 'neuroticism' which has been used in other studies is discussed in the light of the observed negative correlation between N and L.

## I.—INTRODUCTION.

SINCE the publication of preliminary data concerning the J.M.P.I. (Furneaux and Gibson, 1961a), this instrument has been used in research both in Britain and abroad (e.g., Callard and Goodfellow, 1962, Costello and Brachman, 1962). As in all such personality inventories, validity rests upon the supposed veracity of the subject in so far as his test responses correspond with his usual behaviour and his genuine opinions. Research with the adult version, the Maudsley Personality Inventory (M.P.I.), has shown that there is a negative correlation between the N scale and L scale (Gibson 1962a), and furthermore, that groups of subjects with high L scores have demonstrable psychological characteristics in that on other tests, they are significantly different from the rest of the experimental population (Furneaux and Gibson, 1961b, Gibson, 1962b, Furneaux and Lindahl, 1963, Hilgard and Bentler, 1963). In the studies quoted, such subjects have been referred to as 'liars.'

The present study was designed (a) to provide the J.M.P.I. with a lie scale with provisional norms, and (b) to investigate the relationship which such a scale would have with the E and N scales of the Inventory.

## II.—METHOD.

A list of eighteen items was compiled such that frequent endorsement of them would imply an improbably high degree of moral perfection for a child. A number of the items were taken from Hartshorne and May (1928) and suitably anglicized. The phrasing of the items was designed so that endorsement in the 'good' direction necessitates acquiescence in nine items, and denial in the other nine. This is a control against response bias. These eighteen items were incorporated in the standard Inventory.



The combined scale was administered to ninety boys and seventy-one girls in mixed groups. The children were urban dwellers, predominantly working class, and aged 10 to 11 years. The experimenter was careful to assure the children that only he would see their completed sheets and that these would not be shown to their teachers or to anyone else who knew them. This assurance was followed by an appeal for complete truthfulness in their responses. Each item in the Inventory was read out to the children in order to ensure a minimum of mis-reading of the sentences by the less literate of the children.

### III.—RESULTS.

Analysis of the records was done separately for boys and girls. The correlation between each L item and the L scale score attributable to the other seventeen items was calculated. Such a procedure is more accurate for short scales than the usual form of item-scale correlation, although the coefficients obtained must be lower. For no item was there any significant difference between the sexes for the item-scale correlations. The tetrachoric  $r$ 's ranged from .28 to .71; mean = .49. Using all eighteen items, the split-half reliability was found to be .65, using Guttman's formula.

The distribution of data for the boys, girls and total sample, is given in Table 1. The group means for E and N are not significantly different from the norms for the Inventory originally published, nor is there any sex difference for these two scales. The one significant difference in the Table is between the mean L scores of the two sexes ( $t=3.47$  with 159 df,  $p<.001$ ).

TABLE 1  
SCORES OF 161 CHILDREN ON THE THREE SCALES OF THE J.M.P.I.

	n	E		N		L	
		Mean	S.D.	Mean	S.D.	Mean	S.D.
Boys . . . . .	90	12.51	2.72	7.96	3.35	9.06*	4.07
Girls . . . . .	71	12.14	2.61	7.45	3.00	11.21*	3.66
Total . . . . .	161	12.35	2.68	7.73	3.21	10.01	4.04

\* Difference between means on L scale for boys and girls  $p<.001$ .

The product moment correlation between the L scale and the two other scales was calculated separately for boys and girls, as their mean L scores differed significantly. There was no significant correlation between L and E for either sex (boys,  $r = -.09$ ; girls,  $r = -.02$ ), but the correlation between L and N was significantly negative for both sexes (boys,  $r = -.28$ ,  $p<.01$ ; girls,  $r = -.54$ ,  $p<.001$ ).

The distributions of L scores for the two sexes are presented in Table 2 in terms of cumulative percentage frequencies.

### IV.—DISCUSSION.

The split-half reliability of .65 for the L scale is lower than that for the other two scales, but this scale comprises only eighteen items. The coefficient is almost identical in size with that of the adult L scale (.66), which is also of eighteen items.

The rationale of lie scales has been discussed fully elsewhere (Gibson, 1962a), and the present scale was designed to measure the tendency of children to distort their test responses by claiming an unlikely degree of moral perfection. It is postulated that every L score may represent several components; among them are: (a) deliberate misrepresentation of the facts about the known self, and (b) truthful report about habitual good behaviour. Thus, the *unusually* well-behaved child reporting honestly on himself will get a moderately high L score, just as will the naughty child who lies about his bad behaviour. Even so, the content of L items is such that it is hardly possible for a child honestly to endorse all, or nearly all of the eighteen items, however 'good' he may be. In general, virtuous people are restrained from claiming all the virtues by a certain modesty and self-criticism.

It is generally known that girls are better behaved than boys, hence the lower rate of female juvenile delinquency. It was to be expected, therefore, that because the L score partly reflects actual good behaviour, the mean of the girls' group would be significantly higher than that of the boys'.

As has been demonstrated in the studies cited earlier, distributions of L scores obtained with the M.P.I. from various adult populations are characteristically skewed with means low on the scale and a 'tail' of a small proportion of high-scoring subjects. Those adult subjects with scores in the upper tail of the distributions have been referred to as 'liars,' and their psychological characteristics have been the subject of separate study. They have been shown to have N scores significantly lower than the 'honest' majority, but not to differ on E.

With regard to the relationship of L to the other two scales, the present results obtained with children are similar to those obtained with the adult Inventory. There is an important difference, however, with regard to the shape of the distributions of the children's L scores; these have means about the middle of the scale, and a roughly normal shape. Although L correlates negatively with N, and virtually zero with E, there is no indication of a natural cut-off point beyond which subjects may be designated as 'liars.' It is suggested that the tendency deliberately to falsify test responses in order to present a favourable picture also results in lowered N scores, as neuroticism is something which children may be ashamed of: extraversion carries no such stigma.

If children with very high L scores (say, over 1.5 S.D. above the mean of their sex group) are to be regarded as giving somewhat peculiar test responses, the children who have very *low* L scores are also to be regarded as giving somewhat peculiar responses, for a very low L score implies actually boasting about naughtiness. It may be seen in Table 2 that of the boys, 32.6 per cent. (twenty-nine boys) scored lower than 7 on the L scale, but of the girls only 8.4 per cent. (six girls) scored as low. The difference between the two distributions is greatest at this point and it is the contrast in the lower range of the L scale which mainly accounts for the significant difference in means between the sexes. It is the boys who most readily claim to be naughty.

At present the three scales of the J.M.P.I. await proper validation by external criteria. Studies such as that of Callard and Goodfellow (1962) which report small differences in N scores between different socio-economic and educational groups, do not really demonstrate their assumption that 'neuroticism,' in the usually accepted meaning of the term, is associated, positively or negatively, with intelligence or achievement of status.



TABLE 2  
 CUMULATIVE FREQUENCY DISTRIBUTIONS OF L SCORES.  
 Frequencies expressed as percentages of the sex-group.

L Scores . . .	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Boys . . . . .	0	1.1	3.3	11.1	15.5	24.4	32.2	36.6	43.3	52.2	61.1	69.9	78.8	84.4	91.1	93.3	95.6	100	100
Girls . . . . .	0	1.4	1.4	2.8	4.2	7.0	8.4	15.5	26.8	31.0	40.8	49.3	60.5	73.2	80.3	87.3	94.4	95.8	100
Difference .	0	0.3	1.9	8.3	11.3	17.4	23.8*	21.1	16.5	21.2	20.3	20.6	18.3	11.2	10.8	6.0	1.2	4.2	0

\* Divergence between the two distributions is greatest at this point.

It is likely that children with the same N scores but very different L scores cannot be simply equated for 'neuroticism,' because of the significant negative association between these two scales, which reflects attitude to test responses. In the researches with the adult M.P.I. which have been cited, one of the interesting features is that the greatest predictive efficiency of the Inventory concerning external criteria has been obtained by using combinations of the scores on all three scales to get personality profiles for the subjects. The present study has suggested that the L scale applied to children will yield similar results, but that it is more complex in that instead of indicating 'honest' and 'liar' subjects (as in the adult studies) it yields three types, the 'low,' 'average' and 'high' scorers, who may differ in personality characteristics in a meaningful manner.

#### V.—REFERENCES.

- CALLARD, M. P. and GOODFELLOW, C. L. (1962). Neuroticism and extraversion in school-boys as measured by the J.M.P.I. *Brit. J. Educ. Psychol.*, 32, 241-250.
- COSTELLO, C. G., and BRACHMAN, H. M. (1962). Cultural and sex differences in extraversion and neuroticism reflected in responses to a children's personality inventory. *Brit. J. Educ. Psychol.*, 32, 254-257.
- FURNEAUX, W. D., and GIBSON, H. B. (1961a). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.
- FURNEAUX, W. D., and GIBSON, H. B. (1961b). The Maudsley Personality Inventory as a predictor of susceptibility to hypnosis. *Int. J. Clin. Exper. Hypnosis*, 9, 167-177.
- FURNEAUX, W. D., and LINDAHL, L. E. H. (1963). How valid are questionnaire validity studies? (Abstract). *Bull. Brit. Psychol. Soc.*, 16, 19a.
- GIBSON, H. B. (1962a). The lie scale of the Maudsley Personality Inventory. *Acta Psychologica*, 20, 18-23.
- GIBSON, H. B. (1962b). *An Investigation of Personality Variables Associated with Susceptibility to Hypnosis*. Ph.D. Thesis. University of London.
- HARTSHORNE, H., and MAY, M. A. (1928). *Studies in Deceit*. New York: Macmillan.
- HILGARD, E. R., and BENTLER, P. M. (1963). Predicting hypnotizability from the Maudsley Personality Inventory. *Brit. J. Psychol.*, 54, 63-69.

(Manuscript received 4th August, 1963.)



# A STUDY OF STEREOTYPES IN THE CARETAKING OF ENGLISH CHILDREN

BY M. E. HEBRON AND W. LONDON

*(Department of Psychology, University of Hull)*

**SUMMARY.** A questionnaire of stereotypical statements on ways of nurturing children, with each statement scaled along a dimension of detachment/projection was applied to a group of intelligent adolescent girls and their mothers. These results were compared with those for teachers and mothers of pre-adolescent children. The results support the thesis that girls attain stability by their fifteenth year in conflictful areas related to freedom for child and parent. With increasing maturity, conflict occurs in areas less projectively charged. The school would appear to influence most closely the girls' stereotypes regarding discipline, but they reflect the opinions of their mothers most on the protection of children and on the limits of a mother's personal freedom. The transitional age for the attainment of greater independence of attitude in adolescent girls, was located at the fourteenth and fifteenth years.

## I.—INTRODUCTION.

THE major source of adolescent security appears to be the crystallisation of a sense of ego-identity. This turns upon the discovery of appropriate ways for dealing with a variety of situations in the spheres of sexual behaviour, independence from the family, vocational choice and moral values. Girls are more precocious than boys in this, because the roles of mother and wife are typically available to her and are in line with her primary identification. (Mussen, Conger and Kagan, 1963). Girls achieve such stability during the ages 14-16, after which period there is high agreement with statements which express the obvious middle class codes of conduct in stereotyped language requiring little thought or analysis. The extent of the firmness of such beliefs may be measured by the agreement shown by families and teachers (Havighurst and Taba, 1949), but during adolescence such beliefs become more sophisticated and skilfully rationalized (Ausubel, 1954).

Parent-child conflicts tend to occur at the point where the adolescent is discovering appropriate ways of achieving independence from the family, and these fall into two main categories:

- (1) Demand for independence on the part of the child.
- (2) Demand for independence on the part of the mother.

The child's demand for independence creates conflict relating to such matters as the time for getting in at night, use of the family amenities and freedom in choice of friend. Parents demand release from such inconveniences as noise and untidiness, teasing of siblings, silliness and shirking of home duties (Hicks and Hayes, 1938).

## II.—THE PRESENT ENQUIRY.

The purpose of the present enquiry was to examine the following hypotheses:

- (1) By the age 14-16 years, relative stability in ideas concerning the caretaking of children has been reached.
- (2) Most areas of change occur in the areas of freedom for the girl and freedom for the mother.
- (3) Teachers reflect the national stereotypes but evince a bias stemming from their training and role.

## III.—PROCEDURE.

The subjects consisted of 133 girls at a selective grammar school aged 11–18 years, and their respective mothers. For the purposes of this enquiry, the girls were grouped according to age, at yearly intervals from 11 to 15 years, inclusive, but those of 16–18 years comprised one group. These experimental groups were compared with one of forty-one teachers, and ninety-one mothers of pre-adolescent children whose opinions had been obtained in a previous research (Highfield, 1949).

*The questionnaire.* A parental attitude test (Highfield, 1949) consisting of forty-four stereotyped statements on the upbringing of children was used. (See appendix.) The statements refer to five categories of child care, namely, the suppressive, protective, disciplinary, the child's freedom and the mother's freedom. In addition, each statement is scaled along a dimension of projection/detachment on a nine-point scale. Projection is indicated at the low end of the scale and is held to imply emotional involvement when such stereotypes of upbringing are implemented. Thus, the disciplinary items tend to occur in the middle range of the dimension and the suppressive and protective at the low end. The permissive items were rated highly because they were held to be most detached. The questionnaire had been applied during the validation procedure to various groups and the following mean scores for level of detachment/projection was exhibited by each:

26 men and women occupying higher posts of responsibility in education and psychology, most of whom were parents . . . . .	6.3
41 teachers who were not parents . . . . .	5.5
52 fathers of pre-adolescent children . . . . .	4.4
91 mothers of pre-adolescent children . . . . .	3.0
71 mothers of children attending a child guidance clinic . . . . .	3.1

*Procedure.* The adolescent girls were asked to answer the questionnaire under supervision in school, and sealed copies with a covering letter asking them to ensure an independent judgment, were sent to parents with a reply envelope enclosed. The assenting replies to each statement were then collated for each age group. The statements were grouped according to the five areas of child care described above. Then, taking the statements for each area in turn, a rank order was ascribed to it according to the number of assenting replies it received. This was done for each age group and its group of corresponding mothers. Rho coefficients were run between each two sets of ranks for daughters and mothers, and those which failed to reach a 5 per cent. significance level were examined to ascertain the reason for the discrepancy. Where the two rank orders for a statement differed by three places or more, the percentage proportions of parents and girls assenting to it, were tested for significance. Thus the stereotypes on which parents and girls differed most, were drawn from an area of general difference which might be suppressive or permissive and so on. This procedure by which mothers and daughters were compared, was repeated in the examination of results for teachers and girls, for girls and mothers of pre-adolescent children, and for the girls' mothers and the mothers of pre-adolescent children.

## IV.—RESULTS.

(1) *Girls and mothers.*

The correlations for adolescent girls' opinions with those of the two sets of mothers are presented in Table 1 and all four group comparisons summarised in Table 2.



TABLE 1

RHO CORRELATIONS BETWEEN SIX AGE GROUPS OF ADOLESCENT GIRLS AND TWO GROUPS OF MOTHERS IN THE FIVE AREAS OF THE QUESTIONNAIRE.

Age group of adolescent girl	Parent group	Areas of the questionnaire				
		suppressive	protective	disciplinary	child's freedom	mother's freedom
11 years . . . .	own . . . .	79	83	66	80	75
	other . . .	88	62	56	86	(40) 70
12 years . . . .	own . . . .	60	87	84	67	84
	other . . .	63	90	55	92	85
13 years . . . .	own . . . .	(2,8,3) 37	91	69	69	(40) 34
	other . . .	88	83	54	(30) 44 (34)	(40) 61 (39)
14 years . . . .	own . . . .	83	70	(24) 51 (20)	(35) 31 (33)	75
	other . . .	87	85	63	80	72
15 years . . . .	own . . . .	67	88	97	67	98
	other . . .	80	81	70	70	70
16-18 years	own . . . .	84	70	76	(37) 37 (37)	91
	other . . .	80	79	74	(32) 56 (30)	78
All age groups	own . . . .	68	81	74	58	76
	other . . .	81	80	62	71	69

NOTE.—Decimal points omitted.  
The numbers of the items which significantly reduce correspondence are inserted.

TABLE 2

RHO COEFFICIENTS FOR THE FIVE AREAS OF THE PARENTAL ATTITUDE TEST BETWEEN ADOLESCENT GIRLS, TEACHERS AND TWO GROUPS OF MOTHERS.

The groups compared	Areas of the questionnaire				
	Suppression	Protection	Disciplinary	Child's freedom	Mother's freedom
I.—Adolescent girls and their own mothers . . . .	68	81	74	58	76
II.—Adolescent girls and teachers . . . . .	61	66	78	65	72
III.—Adolescent girls and mothers of pre-adolescent children . .	81	80	62	71	69
IV.—Mothers of adolescent girls and mothers of pre-adolescent children . . . . .	78	79	69	72	83

NOTE.—Decimal points omitted.

The size of the correlations, and marked group differences in, and frequency of assent to, some of the statements suggest that the girls tend to differ from their mothers and the mothers of younger children most, during the ages of 13 and 14 years. The chief areas are child's freedom and mother's freedom. The three items listed under suppression and discipline are also calculated to relieve parents of the nuisance value of their daughter's behaviour. All these statements are given below, together with the value ascribed to them on the nine-point scale for detachment projection, and the per centage of girls and mothers assenting to each. After each percentage a plus or minus sign is inserted if it represents an increase or decrease on those of the two previous year groups, with N.C. indicating no change.

<i>Mother's freedom in the area of discipline.</i>		Girls	Own mothers
Item 24 (4.4)	Children should be quiet when father and mother want to rest . . . . .	96 (N.C.)	56.4( -)
Item 20 (3.0)	It spoils a child to consider her whims and fancies . . . . .	63.3( -)	94.0(+)

These differences are significant for the 14-year group and suggest parents are less frequently concerned about Item 24, but more often about Item 20.

<i>Mother's freedom in the suppressive area.</i>		Girls	Own mothers
Item 2 (0.9)	Children should be seen and not heard . . . .	10.0( -)	64.0(+)
Item 8 (2.0)	Contradicting grown-ups is a sign of bad upbringing . . . . .	46.6( -)	72.0(+)
Item 3 (1.0)	Children should obey older people without question . . . . .	73.3(N.C.)	44.0(N.C.)

These differences are significant for the 13-year group and suggest that parents and children differ to an increased extent about Items 2 and 8 which have scores at the projective end of the scale. Item 3 is also highly projective and its presence here indicates that a re-orientation concerning it has decreased the correlation, though the frequency of assent remains unchanged. The change has occurred in the parents, whose weight of assent for this item compared with the others, moves from second place at 11 years to seventh place by 13 years. The girls are more certain that the statement is true than is the case with any of the other suppressive items. It ranks first for them at 11 and 12 years and drops only to second place at 13 years.

<i>Freedom permitted to the mother.</i>		Girls	Mothers own	other
Item 40 (7.5)	A day nursery where mothers can leave their babies during the day is a good thing . . . . .	43.3( -)	84(+)	73
Item 39 (7.5)	Children should decide themselves whether they are to go to Sunday school or not . . . .	66.6(+)	48(+)	26

Both of these items are significant for the 13-year group. Parents show an increased willingness to grant older children the right to decide about Sunday school, thus sacrificing some of their right to freedom from the child's presence about the house, but the girls too show an increase in this attitude. The girls, however, in Item 40 agree less that mothers should be relieved of their caretaking duties. On the part of their mothers there is a greater willingness than is the case with mothers of younger children, to assent to the idea of day nurseries.



<i>Freedom for the child.</i>		Girls		Mothers own other	
Item 30 (6.3)	Children should be allowed to read after they are put to bed . . . . .	13 yrs.	73.3(+)	21	21
		16-18 yrs.	65.0 (NC)		
Item 34 (7.4)	Children should be free to come in and go out of the house as often as they please . . . . .	13 yrs.	33.3 (NC)	64	68
Item 35 (7.4)	Children over 10 years of age should be allowed to go off roaming for the day . . . . .	14 yrs.	76.6(+)	51.7(+)	—
Item 33 (6.7)	If a youngster tumbles, it is best to leave him to pick himself up . . . . .	14 yrs.	80.0(+)	32.9(—)	73
Item 37 (7.7)	Children should be allowed to play in the back streets without supervision . . . . .	16-18 yrs.	75.0(+)	30.8(—)	40
Item 32 (6.4)	Children should be allowed to talk freely at meal times . . . . .	16-18 yrs.	85.0(+)		47

From these indications it would seem that the 13-year group objects to restraint on reading in bed disapproved by the majority of mothers of pre-adolescent children and to a lesser extent by their own mothers. On the other hand, they do not endorse continual coming and going in the house, in spite of the approval of such a state of affairs by two-thirds of the mothers of both groups. This may be because they have been conditioned by reproof about their manner of entering and leaving the house, rather than the frequency.

The 14-year group is concerned more with liberty outside the home rather than with restraint in it. Freedom to injure oneself and to venture out into the unknown receive a swelling volume of assent.

The 16-18-year group pronounce in favour of talking at meals, playing in a safe area without supervision and reading in bed. In all these respects they draw away from their parents and nearer to the greater assurance among the professionally trained caretaker, that these are good things. These viewpoints may be evidence of greater sophistication of judgment in late adolescence among these relatively gifted girls.

All these differences, in the area of the questionnaire concerned with the child's freedom and freedom for the parent, cause the correlation of all age groups with their own parents to be at its lowest.

The results of Table 1 would seem to be in accord with the first and second hypotheses of this enquiry.

(ii) *The influence of family and school.*

Two further comparisons were made by correlating the girls' opinions with those of the teacher sample and the two sets of mothers with one another. The results are given in Table 2, where for convenience, the correlations between the girls and parents of pre-adolescent children, and also their own parents are included.

The most interesting trend of correlation in Table 2 is that of the disciplinary area. Teachers and adolescent girls show as close an affinity about this as do the girls with their own parents, although these particular teachers have in no way contributed to their education. In all other respects the girls' opinions are more similar to that of the parent groups. School exerts a more or less standardised influence concerning achievement and conformity, and thus develops self control through firm insistence from the age of 6 years onwards.

Hence, the high correlation between girls and teachers in this area. In other areas, which are more concerned with face to face relationships in the family circle, teachers' opinions differ from those of girls more, although a considerable likeness remains, especially concerning freedom for adult caretakers. The child's freedom as referred to by the questionnaire, is related to out-of-school activities, but in spite of this, adolescent girls are more in accord with their teachers about permissiveness than they are with their parents. The fact that the correlation between girls and the parents of younger children is higher, suggests that their expectations are still modelled on practices which prevailed when they were younger. This was particularly so in the 11 and 12-year samples of girls, where rho-correlations with the mothers of younger children for child's freedom were as high as .86 and .92.

A high general agreement is shown by the two groups of parents, particularly in the areas indicative of the extremes of projection and detachment. They are least in accord in matters of discipline and freedom for the child, which suggests a change in attitude brought about by a regard for the child's level of competence in self care.

The area of the test in which there is greatest correlation between girls' and parents' opinions is the protective. This is the only area which yields significant correlations throughout. The teachers' deviation from family attitudes, implied in the girls' replies regarding protective caretaking, is marked. It is clear that in suppressive and protective styles girls tend to follow the family pattern, and that the training in child care of the professional educator brings about considerable modification in her attitude.

These considerations have relevance for the third hypothesis. The training of professional educators creates their bias away from naive styles of caretaking grounded in projective attitudes, and towards detached and intellectualised approaches. Their concern with the inculcation of self control would seem to exert the greatest influence on the stereotypical opinions of their pupils. It is through this type of experience that they influence most cultural stereotypes on the caretaking of children.

#### V.—REFERENCES.

- AUSUBEL, H. (1954). *Theory and Problems of Adolescent Development*. Grune and Stratton, N.Y.  
 HAVIGHURST, R. J., and TABA, H. (1949). *Adolescent Character and Personality*. New York: Wiley.  
 HICKS, J. A., and HAYES, M. (1938). Study of the characteristics of 250 junior high school children. *Child Development*, 9, 219-242.  
 HIGHFIELD, M. E. (1949). *The Young School Failure*. Oliver and Boyd.  
 MUSSEN, P. H., CONGER, J. J., and KAGAN, J. (1963). *Child Development and Personality*. Harper and Row, N.Y., 2nd edition.

(Manuscript received 12th November, 1963.)

#### VI.—APPENDIX.

##### A QUESTIONNAIRE FOR ASSESSING PARENTAL ATTITUDE

	Rating for Suppressive (Angry and Rejective):	Rating for Detachment
1. A child should always be prevented from getting its own way with adults . . . . .	.8	.9
2. Children should be seen and not heard . . . . .	1.0	1.0
3. Children should obey older people without question . . . . .	1.0	1.0
4. In a well ordered home a cane is kept handy . . . . .		



5. Children should be punished for all the damage they do .....	1.2
6. Spare the rod and spoil the child is a good saying .....	1.3
7. It is best to smack a child who yells in temper .....	1.4
8. Contradicting grown-ups is a sign of bad upbringing .....	2.0
9. Fidgeting in a child should be stopped .....	2.0

*Protection (Tender and Possessive) :*

10. A child's pleasure should be the first consideration in the home .....	.7
11. Parents always know what is best for their children .....	.9
12. A mother should protect her children whatever they have done .....	.9
13. A child naturally depends on father and mother for his amusement .....	1.3
14. A mother should keep a watchful eye on whatever her child does both inside and outside the home .....	1.4
15. It is dangerous for children under 11 years of age to have pocket knives .....	2.5
16. Parents should make sure that their children are not set on by their play- mates .....	2.7
17. A delicate child needs special protection from other children .....	3.0

*Discipline :*

18. Discipline is more important than freedom .....	2.0
19. Children should be made to help in the house whether they like it or not ...	2.7
20. It spoils a child to consider his whims and fancies .....	3.0
21. A child who owns up to stealing should still be punished .....	3.2
22. Even if children do not like work they should do it .....	3.8
23. A child should have regular jobs which he must do .....	4.2
24. Children should be quiet when father and mother want to rest .....	4.4
25. Boys should do as much housework as girls .....	4.5
26. Children over eight years of age should clean their own shoes .....	4.8
27. Children over eight years of age should make their own beds .....	4.8
28. Regular pocket money should be given weekly .....	5.5

*Child's freedom to experiment :*

29. Children should be allowed to bring others into the house to play .....	5.8
30. Children should be allowed to read after they are put to bed .....	6.3
31. Children should be allowed to talk after they are put to bed .....	6.3
32. It is good to let a child get its own way with other children .....	6.4
33. Children should be allowed to talk freely at mealtimes .....	6.7
34. If a youngster tumbles it is best to leave him to pick himself up .....	7.4
35. Children should be free to come in and go out of the house all day, as often as they please .....	7.4
36. Children over 10 years of age should be allowed to take food with them and go off roaming for the day .....	7.5
37. Children should be allowed to climb high trees .....	7.7
38. Children should be allowed to play in back streets without supervision ....	7.7

*Mother's freedom :*

39. Children should be allowed to bring others into the house to play ..	7.3
40. It is good when children can get their mid-day meal at a school canteen ..	7.5
41. Children should decide themselves whether or not they go to Sunday School	7.5
42. A day nursery where mothers can leave their babies during the day is a good thing .....	8.0
43. Children would rather go out with their friends than with their parents ..	8.5
44. Children should start school as early as two years of age .....	8.5
45. A mother should find her pleasures outside the home .....	8.5
46. A family is happiest when the mother as well as the father goes out to work	8.5

# AN INVESTIGATION INTO THE COMPARATIVE RESPONSE OF BOYS AND GIRLS TO SCRIPTURE AS A SCHOOL SUBJECT IN CERTAIN CO-EDUCATIONAL GRAMMAR SCHOOLS IN INDUSTRIAL SOUTH WALES

BY R. R. DALE AND J. A. JONES

(*University College of Swansea*)

**SUMMARY.** An attitude scale and questionnaire were administered to pupils in the fourth forms of ten co-educational grammar schools in industrial South Wales, in order to compare the response of the sexes towards Scripture. The attitude of both sexes was favourable, the girls being more favourable than the boys. The scores for various sections of the syllabus were examined. 'The Acts' were liked least by both sexes. Interesting social class differences emerged. Starting with Social Class I, girls' responses became progressively *more* favourable and boys' responses *less* favourable, in successive social classes. The sex difference, therefore, tended to become progressively greater. Strong attachment to Sunday School and/or Place of Worship was associated with more favourable response, and less difference both between the sexes and the social classes. *In this sample* attitudes were slightly more favourable when pupils were taught by a mistress.

## I.—INTRODUCTION.

THIS investigation is concerned with differences between girls and boys in the fourth forms of certain co-educational grammar schools, with regard to their respective responses to Scripture. Their responses are interpreted as being an expression of their attitude—interest reaction to Scripture as a school subject. A comparison is made of the differences in response to various sections of the subject and the hypothesis is examined that the different social class backgrounds would be associated with changes in response.

## II.—OUTLINE OF PREVIOUS RESEARCH.

Previous research relating to this subject is not extensive. Lewis (1913), investigating the popularity and unpopularity of school subjects in elementary schools, found that Scripture tended to be placed low in the list. Shakespeare (1936) came to the same conclusion, as did Houslop and Weeks (1948), though their enquiry was among boys in a grammar school. In the last of these researches there was a tendency for the subject to be generally disliked. On the other hand, Moreton (1944), Glassey (1945) and Daines (1949) found the attitude of adolescent boys and girls to *religion* to be moderately favourable. In Glassey's enquiry, the attitude to religion became less favourable as the child moved up the school.

Lewis found that Scripture tended to be more popular with girls than with boys. So did Shakespeare. Glassey found that the attitude of grammar school girls was decidedly more favourable than that of the boys. The work of Daines produced no evidence of any clearly marked *religious* leanings in adolescent girls as opposed to boys, but he considered that possibly the girls showed a more pronounced interest. Forrester (1946) discovered only a small number of differences between the attitudes of adolescent girls and boys, and when they did occur they were seldom very great.



When girls and boys are taught together, it is important for their teachers to know the extent of any differences in their interest in, and attitude towards, Scripture and its various sections, and also whether any differences are of practical importance in the teaching of the subject. Previous research suggests that some differences do exist. Where some indication of the nature and extent of these differences can be given, an appropriate allowance can be made for these in teaching, so that the respective interests of the girls and boys can be fully utilised in building up the right attitude to the subject.

### III.—PROCEDURE.

Twenty-one (out of a total of thirty-three) co-educational grammar schools in industrial South Wales were visited, and their co-operation sought in carrying out this investigation. Sixteen schools willingly provided their Scripture scheme of work.

On the basis of information received from thirteen schools, and using also The Syllabus of Religious Instruction for the Schools of Wales, Section B of the questionnaire was prepared. (See Appendix). This contained a list of twelve topics for Scripture lessons, classified under three headings: The Old Testament, The Gospels, The Acts of the Apostles. For each topic, pupils were to underline one of five responses: Like very much, Like, Undecided, Dislike, Dislike very much. The Likert scoring technique (1932) was adopted, whereby values of 1 to 5 were assigned to the five possible responses, the highest value being given to 'Like very much.'

When Sections A, B and C of the questionnaire were complete, a pretest was carried out with one form of grammar school pupils to discover whether there were any difficulties and also to obtain an estimate of the time taken to answer these sections. Some changes were subsequently made in the wording of the questionnaire.

In the preparation of the Attitude Scale in Section D (see Appendix), Edwards and Kilpatrick's method (1948) was followed up to a point. Statements were collected, judged and scaled according to the Thurstone procedure (1929). The Likert method of scoring (1934) was used, and the criterion of internal consistency was applied to eliminate undifferentiating statements.

Two-hundred statements, collected from various sources, were subjected to the criteria suggested by Wang (1932) and Edwards (1957), for writing attitude statements. Fifty-eight statements remained. The evaluation of these by thirty-eight judges formed the basis of the assignment of a scale value and a Q value to each statement.

All statements with a higher Q value than the original average Q value (1.33) were eliminated. Thirty-three statements remained. These were re-examined in the light of the interpretation of answers (McNemar, 1946). This reduced the number to twelve. One, although considered unsuitable, was retained for checking purposes.

In order to increase the number of statements, especially in the centre groups, some of the statements with a higher Q value than the average were re-examined. Four were added to the list. These sixteen statements yielded an average Q value of 1.15 and the Q value of no statement was more than 2.0 (Vernon, 1953).

In the scoring of this scale, the scale values of the statements must be taken into account when assigning weights to the responses (Strongly Agree, Undecided, Disagree, Strongly Disagree). In the 'Thurstone Weighted Method' of Likert, Roslow and Murphy (1934) the values assigned are based upon the scale value of the Thurstone Method. (Butcher, 1956.) Beaumont (1957) used a modified



Thurstone Weighted Method which simplified the scoring. Scale values, reduced to the nearest whole number, became the 'Strongly Agree' alternative for each statement. The high scores were always assigned to the favourable end of the scale. This modified method was adopted here.

The scale of sixteen statements was then administered to a pretest group of sixty-two pupils, mainly to discover undifferentiating statements and to obtain a measure of the reliability of the scale.

For item analysis, the scores of twelve pupils in the upper half and twelve pupils in the lower half were used. Statements were to be rejected if they failed to reach a certain significant figure. (Beaumont, 1957.) For only one statement was the difference between the upper and lower halves less than the corresponding significant figure; the statement was accordingly eliminated. This was in line with expectations—this was the statement that had been retained for checking purposes.

The reliability of the scale was calculated using the split-half method. The coefficient of correlation between the X and Y halves of the scale was +.973.

Many of the questions in the questionnaire section were open-ended, as it was thought that the comments would increase the value of the results; this part of the work is too lengthy for inclusion here and will be dealt with elsewhere.

The questionnaire and attitude scale were personally administered in ten schools. No member of the school staff was ever present while the pupils were writing their answers. Each pupil was given the sheets of questions and a copy of the instructions. The latter was read to them, reference being made to the questions wherever necessary. The purpose of the investigation was explained. The pupils were asked to be as frank as possible and were reassured that all answer papers would remain anonymous. The necessity for them to think carefully before answering every question was emphasised. They were invited to ask questions and quite a number did so.

The tested population consisted of 172 girls and 108 boys, the average age being 15 years 2 months, and 15 years 1 month, respectively. The majority of both girls and boys were non-Welsh speaking. The pupils were classified into four social groups, according to father's occupation. These groups were: (1) Higher non-manual; (2) Lower non-manual; (3) Skilled manual; (4) Semi-skilled and Unskilled manual. (Glass, 1954, and Floud, Halsey and Martin, 1956.) In five of the schools the pupils were being taught Scripture by a mistress, and in the remaining five by a master. 66.8 per cent. of the girls were regular attenders at Sunday School, but only 43.9 per cent. of the boys. 69.6 per cent. of the girls were regular attenders at a place of worship, but only 47.2 per cent. of the boys.

A comment is needed on the disproportion between boys and girls in the sample. In order to obtain the willing co-operation of the school and forestall any deliberate selection of pupils it was decided to administer the questionnaire to one entire form in each school. To secure a reasonable homogeneity of age and ability only fourth 'A' streams were chosen. To the surprise of the experimenters this resulted in an imbalance of the sexes, and the question arose whether the two groups were sufficiently representative. Inquiries in the schools produced the following evidence. The pupils in the tested forms were all those in the stream who took 'statutory scripture,' except for thirty-four absentees, of whom eighteen were girls and sixteen boys. The latter fact reduced the discrepancy to a three to two proportion. None of the forms consisted entirely of pupils who intended to take Scripture in the Ordinary Level examination, so that the sex imbalance was not produced in this way. In five of the schools selection for the 'A' stream was on merit only; in two of these cases (Schools



3 and 10) the proportion of girls to boys in the school was three to two, in two others (Schools 2 and 7) there were more girls than boys in the school, School 2 having 270 girls to 230 boys; in the remaining school (9) the sexes were equal in number. In School 4 the pupils in the 'A' stream took no Latin, in School 8 they were the pupils who opted to take Latin. In School 5 allocation was by age only. In School 1 the 'A' stream was one of two parallel Arts forms, and the Science stream had thirty-one boys and four girls. In School 6 there were fairly equal numbers of boys and girls in the school, and the only other information available was that 'the fourth forms were considered equal forms.' This form had seventeen girls and twelve boys.

Though the overall appraisal of the sample is reassuring, with no 'special selection' in six of the schools and the Latin and non-Latin streams balancing each other in two more, there remain two schools (1 and 6) in which there may have been some special selection. Here, although the boys and girls may once again be said to balance each other because they both may have opted for an Arts course, yet in another sense, the small proportion of boys in the Arts class might produce a result which is not as representative of their sex as the corresponding score is with the girls. In the analysis of the results, it was, therefore, necessary to keep this difficulty in mind: the principal corrective was that the researchers gave careful attention to the pattern produced by the results from the ten individual schools, and to the effects produced on the findings by the exclusion of Schools 1 and 6.

#### IV.—RESULTS.

The response scores which represent the pupils' attitude to Scripture lessons in general are given in Table 1.

TABLE 1  
MEAN SCORES OF THE GIRLS' AND BOYS' GROUPS IN EACH OF THE TEN SCHOOLS, REPRESENTING THEIR ATTITUDE TO SCRIPTURE LESSONS IN GENERAL.†

Schools	No. of pupils		Section B			Section D		
			Possible range 12-60			Possible range 29-121		
	Girls	Boys	Girls	Boys	Difference	Girls	Boys	Difference
1	18	8	49.4	46.0	3.4	104.9	97.3	7.6
2	18	12	46.6	43.5	3.1*	90.7	72.8	17.9†
3	19	9	48.7	40.8	7.9†	99.2	91.1	8.1
4	13	6	45.9	46.8	-0.9	93.0	101.7	-8.7
5	17	15	50.7	44.1	6.6†	106.4	81.8	24.6†
6	17	12	45.9	48.3	-2.4	99.8	104.5	-4.7
7	16	10	49.7	45.2	4.5	100.7	88.1	12.6
8	16	10	48.5	43.8	4.7*	101.8	91.6	10.2*
9	22	14	44.7	41.5	3.2	88.0	72.8	15.2
10	16	12	48.2	47.0	1.2	96.0	89.7	6.3
Total . . . .	172	108	—	—	—	—	—	—
Mean . . . .			47.8	44.6	3.2†	97.9	87.6	10.3†

\* Significant at the 5 per cent. level.

† Section B is the Questionnaire, Section D the Attitude Scale.

‡ Significant at the 1 per cent. level.

The *girls'* mean scores indicate that they like Scripture as a school subject (Section B), and give a favourable response to it on the Attitude Scale (i.e., Section D). The exclusion of schools 1 and 6 made no difference to the findings, even though the attitude score was thereby slightly reduced (to 96.7).

The *boys'* mean scores indicate that they have a moderate liking for Scripture (Section B), and give a moderately favourable response to it on the Attitude Scale. As in the case of the girls, the exclusion of schools 1 and 6 reduced the attitude score (to 84.4); their score in Section B was also slightly lower. The findings are unchanged.

The difference between the girls' and the boys' mean scores (3.2 in a range of forty-eight for Section B, and 10.3 in a range of ninety-two for the Attitude Scale) is statistically significant at the 1 per cent. level of confidence in both Sections. If schools 1 and 6 are omitted, the difference is increased to 12.3. It would thus appear that the girls are rather more interested than the boys in Scripture as a school subject. This result was checked by the use of a small sample statistical technique whereby the school means were used as the unit instead of the scores of individuals. This method enabled us to see whether the difference between the response of the boys and the girls was found sufficiently consistently in the different schools. The statistical outcome was the same as in the case of the previous tests. For Section B and for Section D the difference was in the expected direction in each of eight schools, and in the reverse direction in the two others. The Wilcoxon sign-rank non-parametric test again confirmed these differences when the boys' and girls' means within each school were taken as the unit.

As one would expect, there is a considerable overlap between the scores of boys and girls. In Section B (Questionnaire) 32 per cent. of the boys had a higher score than the mean score of the girls, while 21 per cent. of the girls had a lower score than the boys' mean score. Similarly, in Section D (Attitude Scale) 35 per cent. of the boys had a higher score than the mean score of the girls, and 19 per cent. of the girls scored less than the boys' mean score. An interesting point is that in both the Questionnaire and the Attitude Scale, there was a very slightly greater percentage of boys than of girls in the highest scoring group. This difference was not statistically significant and might be due to the slight element of special selection in the boys' sample.

Analysis of the Questionnaire scores for each of the three sections of the Scripture Syllabus is given in Table 2.

The table shows that neither the girls nor the boys like each section equally well. Both the girls and the boys like The Gospels most. (Girls have a great liking for it, boys like it.) Next, for both the girls and the boys, comes The Old Testament. (Girls and boys like it.) Least liked by both the girls and the boys is The Acts of the Apostles. (Girls have a moderate liking for it, boys are undecided but tend towards having a moderate liking.)

The difference between the responses of the sexes is not the same in each section of the syllabus. For The Gospels, the difference between their mean scores (10 per cent.) is statistically significant at the 1 per cent. level of confidence. The direction of the difference is the same in all ten schools and the means of nine girls' groups are higher than that of any of the boys' groups. In the other two sections, the differences between the girls' and the boys' mean scores (1.5 per cent. for The Old Testament and 4.5 per cent. for The Acts) are statistically insignificant. In each section of the syllabus, the girls' mean score is higher than that of the boys. For The Old Testament and The Acts, the girls' mean scores are lower than those of the boys in some schools, higher in others.



TABLE 2

THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN THE GIRLS' AND THE BOYS' MEAN SCORES FOR THE THREE SECTIONS OF THE SYLLABUS IN EACH OF THE TEN SCHOOLS.

School	No. of Pupils	Mean Scores		
		Old Testament	Gospels	Acts
1	Girls 18	15.7	18.9	14.8
	Boys 8	15.9	15.3†	14.9
2	Girls 18	15.3	17.9	13.4
	Boys 12	15.3	16.3	11.8
3	Girls 19	16.6	17.9	14.3*
	Boys 9	16.1	13.4†	11.3
4	Girls 13	15.2	18.0	12.8
	Boys 6	16.8	16.8	13.2
5	Girls 17	17.0	18.7	15.1*
	Boys 15	15.2	16.5	12.5
6	Girls 17	14.8	17.7	13.5
	Boys 12	15.8	17.4	15.2
7	Girls 16	16.8	18.3	14.6
	Boys 10	15.7	16.4	13.1
8	Girls 16	15.8	18.3*	14.5
	Boys 10	16.0	15.6	12.2
9	Girls 22	16.2	16.9	11.6
	Boys 14	14.3	14.9	12.4
10	Girls 16	15.4	18.2	14.7
	Boys 12	16.2	17.2	13.7
Total .....	Girls 172	—	—	—
	Boys 108			

\* Significant at the 5 per cent. level.

† Significant at the 1 per cent. level.

It would thus appear that so far as The Old Testament and The Acts are concerned, the difference between the girls' and the boys' responses are not sufficiently great to be of practical importance in the teaching of Scripture in co-educational grammar schools. The difference between their responses to The Gospels is greater and may be of practical importance, although it should be remembered that even here their responses are more alike than unlike, and that even the boys claim to like this section of Scripture; their 'popularity' score for it *in this section* places it higher than the other two sections.

In another part of the Questionnaire, as a partial check on the above results, the girls and boys are asked to indicate their order of preference for the three sections of the syllabus. A difference between their preferences then emerges. The girls' order of preference is: (1) The Gospels, (2) The Old Testament, (3) The Acts. This order is the same as before. The boys' order of preference

is: (1) The Old Testament, (2) The Gospels, (3) The Acts. This order differs from that found previously, in that the first two sections of the syllabus have changed places. This shows the value of the inclusion of check questions. One reason for the reversal of the rank order of the two sections is that in Section B of the enquiry, the boys placed them very close in 'popularity,' averaging 15.6 for The Old Testament and 16.0 for The Gospels, out of a possible 20. The emphasis here is on similarity rather than difference. As, however, the difference becomes more substantial for the check question, there must be one or more additional reasons for the change in rank. The fact that 'The Old Testament' was printed on the first line, with the two other titles on successive lines below might have increased its score slightly, but this alone would probably be too small an influence. Two other suggestions are made. First, that the thirty-two topics given in each of the relevant sections of the questionnaire as examples of the Scripture of The Old Testament and The Gospels, taught in the sample schools, might not have been quite as representative as we had hoped, in spite of the very great care that was taken. Second, that some attitude or prejudice may have affected the response of the boys to either of the sections as a whole which had little effect on their response to individual topics within those sections. To keep this matter in the right perspective, it should be mentioned that if we weight these preferences 1, 2 and 3 for first, second and third preferences, the scores are 1.71 (The Old Testament) and 1.97 (The Gospels). The change in attitude score which produces the reversed rank order is not great enough to cause serious concern when placed beside the cases where there is agreement, but it is sufficient to require comment.

When the scores are analysed with regard to pupils' language differences, no conclusion can be drawn about the Welsh-speaking group as a whole, because of the nature of the sample, which is mainly drawn from one school. A little can be said, however, about the difference between the attitudes of the sexes. In both the Welsh-speaking and non-Welsh-speaking groups, the girls' mean scores (Sections B and D) are higher than those of the boys. In the Welsh-speaking group, the difference between mean scores is statistically insignificant, perhaps because of the smallness of the tested population, but in the non-Welsh-speaking group the difference is statistically significant at the 1 per cent. level of confidence.

Certain trends are noticeable when the differences between the sexes are examined in relation to social class. Starting from fairly similar responses of boys and girls (Sections B and D) in Social Group 1, the responses of the non-Welsh-speaking *girls* show a tendency to become progressively *more favourable* in successive social groups, whereas the responses of the non-Welsh-speaking *boys* show a slight tendency to become progressively *less favourable* in successive social groups. In the case of the girls, the tendency is surprisingly regular from Group 1 to Group 4, but in that of the boys it only consists of a more favourable response in Group 1 than in all other Groups, the scores of these latter groups differing very little between themselves. The above social class differences were not due to between school differences. The difference between the mean responses of the sexes, therefore, tends to become progressively greater in successive social groups.

In Social Group 1, the difference between the sexes is statistically insignificant. In Group 2, there is a difference in the results for Sections B and D, the difference between mean scores being statistically significant in Section D only. The difference between mean scores is statistically significant at the 1 per cent. level of confidence in Group 3, and at least at the 5 per cent. level in Group 4.



It would thus appear that a change in social group is accompanied by a change in the responses of both the girls and the boys, and also in the amount of the difference between them. This seems to support one of Forrester's conclusions (1946) that one of the chief reasons for differences in attitudes may be personality differences which go back to constitutional or environmental differences. She says that the question then is what kind of home life produces the most positive attitude to the topic under investigation. In this enquiry, the only indication of home life is that provided by the social group to which a pupil belongs. So far as the girls are concerned, the most positive response is given by those who belong to social group 4, and the least positive response by those in Group 1. So far as the boys are concerned, the most positive response is given by those in Group 1, the least positive by those in Group 4.

The mean responses of the non-Welsh-speaking girls are more favourable than those of the non-Welsh-speaking boys, whether they are taught by mistress or master, the difference between mean scores being statistically significant at the 1 per cent. level of confidence (Sections B and D) when the teachers are mistresses, and at the 1 per cent. level in Section B and 5 per cent. level in Section D when the teachers are masters. Nevertheless, a difference in the sex of the teacher is accompanied by a difference in both the girls' and the boys' responses, as follows. Both the girls and the boys tend to give slightly more favourable responses when they are taught by a mistress. The difference between the responses of the girls and the boys taught by mistresses tends to be less than that of the girls and the boys taught by masters. As there are only five masters and five mistresses these particular results are not presented as findings, but await the verdict of future research.

Although the samples are small, the above tendencies are also observed within each social group. Furthermore, it appears that the observations made previously regarding social group differences, apply whether the pupils are taught by mistresses or by masters.

The results when scores are grouped according to attendance at Sunday School are fairly similar to those when scores are grouped according to attendance at a place of worship. The mean scores of the non-Welsh-speaking girls are higher than those of the non-Welsh-speaking boys in both cases. Certain other tendencies are observable. The responses of the girls whose attachment to a religious institution is strong (i.e., regular attendance) tend, as one would expect, to be more favourable than those of the girls whose attachment is weak or non-existent. The same is true also for the boys, for whom this tendency is more pronounced. The difference between the responses of the girls and boys whose attachment to a religious institution is strong, is less than that between the girls and boys whose attachment is weak or non-existent.

The pattern of the results indicates that these observations probably apply to each social group, though here the samples are small.

Where there is a strong attachment to a religious institution, the tendencies of the girls' and the boys' responses to vary according to their social group are not much in evidence. However, where there is a weak or non-existent attachment to a religious institution, the tendencies of the girls' and the boys' scores to vary according to their social group are clearly in evidence.

#### V.—COMMENT.

Much of this enquiry has been concerned with an examination of differences in the response of girls and boys towards Scripture. The extent of the agreement is, however, more striking and more important. In Section B, where the possible score range is 48, for attitude to the subject as a whole, the boy-girl

difference in score in eight out of the ten schools is 4.7 or less. Similarly, in the attitude scale, where the maximum possible score is 121, with a range of 92, the girls' score is 98, while that of the boys is 88. In no overall or section score does one sex record liking and the other sex dislike. The study brings out clearly the difference between something which is statistically significant and something which is of practical significance. A number of the differences reported in the study have statistical significance, but none of them would appear to be of serious practical significance for class organisation, though a knowledge of them might well be of value to a teacher of Scripture.

The size of the sample, ten schools and 280 pupils, was limited by practical considerations, as was its restriction to one area. For this reason, we present these findings tentatively, hoping that they will form a useful guide for future hypotheses and research. Of particular interest in this respect are the results with regard to social class.

## VI.—APPENDIX.

### ABBREVIATED QUESTIONNAIRE.

#### SECTION B.—TOPICS FOR SCRIPTURE LESSONS.

##### THE OLD TESTAMENT : THE HISTORY AND RELIGION OF ISRAEL.

- (1) *The history of the Israelites from Abram to Joseph.*  
e.g., God's covenant with Abram, Isaac blessing Jacob, Jacob's dreams, Joseph's dreams, Joseph sold into slavery by his brothers, Joseph in Egypt.  
Like very much / Like / Undecided / Dislike / Dislike very much.
- (2) *The history of the Israelites in Egypt, the Exodus and Wanderings.*  
(Examples omitted.)  
Like very much, etc.
- (3) *The history of the conquest and settlement of Canaan.*
- (4) *The history of the United Monarchy.*

##### THE GOSPELS : THE LIFE AND TEACHING OF JESUS CHRIST.

- (5) *Stories of the birth of John the Baptist, and of the birth and childhood of Jesus.*
- (6) *The teaching of Jesus.*
- (7) *The Miracles of Jesus.*
- (8) *The events of the last week of Jesus' life.*

##### THE ACTS OF THE APOSTLES : THE BEGINNING OF THE CHRISTIAN CHURCH.

- (9) *The promise and the coming of the Holy Spirit.*
- (10) *The Church in Jerusalem.*
- (11) *The Church in Samaria.*
- (12) *The Church among the Gentiles.*

#### SECTION C.

- (1) What do you think of Scripture as a school subject?  
Like very much / Like / Undecided / Dislike / Dislike very much.  
Give a short reason for your answer



- (2) Number the following sections of Scripture lessons in your order of preference :  
*The Old Testament*  
*The Gospels*  
*The Acts of the Apostles*
- (3) Give a short reason why in question 2 you like your first choice best and your last choice least.  
Reason for my first choice is  
Reason for my last choice is
- (4) Where would you place Scripture in a list of your subjects in order of preference? Underline your answer.  
Very High / High / Middle / Low / Very Low.
- (5) What do you like most in Scripture lessons?
- (6) What do you dislike most in Scripture lessons?
- (7) Do you intend taking Scripture as one of your examination subjects at the Ordinary Level?  
Yes / Undecided / No.

## SECTION D (Attitude Scale)

- (1) We ought to study Scripture carefully and thoroughly.  
Strongly Agree / Agree / Undecided / Disagree / Strongly Disagree.
- (2) Scripture lessons do not deserve the praise that some pupils give them.  
Strongly Agree, etc.
- (3) Scripture lessons are important for all pupils.
- (4) I have no desire at all to learn Scripture.
- (5) I would not encourage anyone to study Scripture seriously.
- (6) Without Scripture lessons our education would be incomplete.
- (7) I would never study Scripture in school unless I had to.
- (8) Scripture is a very interesting subject.
- (9) Scripture lessons are a waste of time.
- (10) I think that when pupils speak ill of Scripture lessons they are right.
- (11) Scripture lessons are a necessary basis for the best way of life.
- (12) What I learn in Scripture lessons is usually worthwhile.
- (13) School would be better off without Scripture lessons.
- (14) More time should be given to Scripture lessons.
- (15) I should like to stop Scripture from being taught in schools.

## VII.—REFERENCES.

- BEAUMONT, G. P. (1957). *A Study in the Construction and Application of a Scale Measuring the Attitude of Adolescents towards the Use of Force*. M.A. Thesis, London.
- BUTCHER, H. J. (1956). A note on the Scale Product and related methods of scoring attitude scales. *Brit. J. Psychol.*, 47.
- DAINES, J. W. (1949). *A Psychological Study of the Attitude of Adolescents to Religion and Religious Instruction*. Ph.D. Thesis, London.
- EDWARDS, A. L., and KILPATRICK, F. P. (1948). A technique for the construction of Attitude Scales. *J. Applied Psychol.*, Vol. 32, No. 4.
- EDWARDS, A. L. (1957). *Techniques of Attitude Scale Construction*. Appleton-Century-Crofts, Inc., N.Y.
- FLOUD, J. E. (Ed.), HALSEY, A. H., and MARTIN, F. M. (1956). *Social Class and Educational Opportunity*. Heinemann, p.150.

- FORRESTER, J. F. (1946). *A Study of the Attitudes of Adolescents to their own Intellectual, Social and Spiritual Development*. Ph.D. Thesis, London.
- GLASS, D. V. (Ed.) (1954). *Social Mobility in Great Britain*. Routledge and Kegan Paul. Chapter 12.
- GLASSEY, W. (1945). The attitude of grammar school pupils and their parents to education, religion, and sport. *Brit. J. Educ. Psychol.*, Vol. 15.
- HOUSLOP, H. L., and WEEKS, E. J. (June, 1948). The interests of school children. *The School Science Review*.
- LEWIS, E. O. (June, 1913). Popular and unpopular school subjects. *J. Exper. Ped.*
- LIKERT, R. (1932). Technique for the measurement of attitudes. *Arch. Psychol.*, Vol. 22, No. 140.
- LIKERT, R., ROSLOW, S., and MURPHY, G. (1934). A simple and reliable method of scoring the Thurstone Attitude Scales. *J. Soc. Psychol.*, No. 5.
- MCNEMAR, Q. (1946). Opinion-attitude methodology. *Psychol. Bull.*, 43, p. 317.
- MORETON, F. E. (1944). Attitudes to religion among adolescents and adults. *Brit. J. Educ. Psychol.*, Vol. 14.
- PRITCHARD, R. A. (1935). The relative popularity of secondary school subjects at various ages. *Brit. J. Educ. Psychol.*, 5.
- SHAKESPEARE, J.-J. (1936). An enquiry into the relative popularity of school subjects in elementary schools. *Brit. J. Educ. Psychol.*, 6.
- THURSTONE, L. L., and CHAVE, E. J. (1929). *The Measurement of Attitude*. Univ. of Chicago Press.
- VERNON, P. E. (1953). *Personality Tests and Assessments*. Methuen, p. 152.
- WANG, C. K. A. (1932). Suggested criteria for writing attitude statements. *J. Soc. Psychol.*, 3.
- A Syllabus of Religious Instruction in the Schools of Wales* (1945). Univ. of Wales Press Board, Cardiff.

(Manuscript received about 18 months ago.)



# WRITTEN COMPOSITION AT 10 YEARS AS AN ASPECT OF LINGUISTIC DEVELOPMENT

## A LONGITUDINAL STUDY CONTINUED

BY OLIVE C. SAMPSON

(*Department of Education, University of Manchester*)

SUMMARY. The fourth instalment of a longitudinal study of linguistic development is reported. The current research concerns the children's skill in written composition at 10 years of age. Their work is assessed in various ways (by the 'subordination index,' vocabulary studies, and impression ratings). The results are correlated with other linguistic measures currently or previously established, and with non-verbal intelligence. The children's adjustment is also considered.

### I.—INTRODUCTION.

THIS study continues the work reported in 1956, 1959 and 1962 on the linguistic development of a sample of fifty children representative of all levels in the community. The first of these researches described their speech development from 18-30 months of age, the second dealt with speech and language at 5, the third considered their reading skill at 8, in relation to linguistic and other factors. The present deals with their written work at the age of 10+, as an aspect of total linguistic development. The work of their classmates was also reviewed for the general information it provides on the skill in composition of the age-group in question, but is not here reported on.

### II.—METHOD.

Contact was made with the children of the sample who were still available (forty-five), both at home and in their twenty-eight schools. The following material was obtained:

(1) *Two twenty-minute compositions*, one on a factual, the other on an imaginative topic. These were written on separate occasions under the class teachers in normal classroom conditions.

(2) Results of a *reading test* (Holborn Reading Scale) and of

(3) A *vocabulary test* (from Stanford-Binet Intelligence Scale Form L-M). These tests were done by the sample children in individual interviews with the writer.

(4) *The Bristol Guide (The child in school)*. This was completed by the class teachers in respect of the sample children.

(5) A *visit* by the writer to the homes of the sample brought background information up to date.

### III.—APPRAISING CHILDREN'S WRITTEN COMPOSITIONS.

It was desired to ascertain both the characteristics and general level of skill found in 10-year-olds' compositions and also to see how far the ability of the sample children accorded with their present and past linguistic development in other fields. The differential influences of sex and subject matter were also to be considered.

The appraisal of written composition is a complex task. The literary approach is highly subjective, and relevant scientific studies in this field are scarce, though suggestive. They can be classed under three heads, according to their main point of view, whether this is developmental, verbal, or from the examiner's angle. All these have been laid under contribution in the present research. On the evolution of written from oral language, O'Shea (1907) and Labrant (1933), for example, make different but invaluable contributions from the developmental angle. O'Shea's *Language Development and Education* (1907) is an early and excellent example of its sort. Starting with studies of his own children's verbal performances in infancy, he also quotes other authorities and leads on to a consideration of written composition and the principles of language learning. O'Shea's book is stimulating and comprehensive. It anticipates much modern thinking. More recent works, such as LaBrant's *A Study of Certain Language Developments of Children in Grades Four to Twelve, inclusive* (1933) concentrates on a narrower area and is far more statistically sophisticated. Basing her approach on that of M. E. Smith (1927) and M. M. Nice (1925), when dealing with oral language, she concludes, after comparisons with some other measures, that the subordination index is the best measure of maturity in written composition. C. T. Ford, in New Zealand, in *Development in Written Composition During the Primary School Period* (1954), deals with the age range 7-15 in his qualitative and quantitative analysis. He finds a positive correlation between composition writing skill and I.Q. but "discernable improvement in quality was comparatively rare from one age group to the next." Girls were superior in 'observed differences.' Earlier (1930), Bushnell had had interesting results comparing the oral and written work of tenth grade children by ingenious procedures. Thought content, sentence structure, errors and vocabulary were studied. Written composition was found to be better all round but not essentially different from oral work.

Perhaps the most notable contribution to the verbal analysis of composition is provided by J. W. Chotlos in *A Statistical Analysis of Individual Written Language Samples* (1944). He uses the type-token method described by Johnson (1944) by which the numbers of different words in any grammatical category (type) is related to the total number of words (tokens) in the sample. He correlates these with I.Q. and C.A. Though he is mainly dealing with children of 12-15, his results are of interest. He finds that the higher the I.Q. and the higher the age-level, the more highly differentiated is the language structure of the writers. But "on the assumption that the correlation of C.A. and I.Q. is zero, the multiple correlation co-efficient of the type-token ratio with C.A. and I.Q. is only .608, a result which suggests that C.A. and I.Q. are not sufficient factors for completely determining the type-token ratio." His method has been (in part) adopted in the present research.

Composition skill is also sometimes approached from the point of view of the examiner and work from this angle had to be considered in making a choice here. Some advocate tests. For example, Morgan (1956) investigating 'Verbal abilities in Primary School children' developed a number of tests. A prominent general verbal factor 'literacy' was found. Various fluency measures in particular had a high 'literacy' loading. Other writers debate the pros and cons of error-analysis *versus* impression marking. Penfold (1956) shows the great variation between examiners using the analytic method, but in the previous year (1955), Nisbet had, nevertheless, found that both methods give the similar results. Wiseman (1949, 1956) prefers the impression method for statistical and practical reasons. On the whole, the balance seems to favour impression marking and this was adopted.



## IV.—THE COMPOSITIONS.

In obtaining examples of children's written work, they may be given minimal directions, as Chotlos chose to do: "You are to write about anything you want to write about. Just make it up as you go along." Alternatively, definite subjects may be set, as in the sixty-eight topics of the Scottish Council for Research in Education's (1961) spelling survey. In the present case, in spite of Bilbrough's (1955) rather negative findings in connection with titles, a choice of topics of two contrasted types, factual and imaginative, were offered and short compositions (twenty minutes' writing time) were asked for on each. Predicting response to titles is feasible up to a point (Bilbrough, 1955) and the children's choices illustrated this. The factual titles (Set A) were as follows:

- (1) Do you like cooking? Tell how to cook something.
- (2) Looking after a pet.
- (3) A ride in a bus or train.
- (4) My best toys.

The imaginative titles (Set B) were:

- (5) Black, white and grey.
- (6) A dream.
- (7) A letter from a planet.
- (8) A talk with an old, old man.
- (9) A voyage to a fairy island.

The children of the sample and their classmates wrote according to their choice in the conditions already mentioned. The results of the sample children are examined below in some detail.

TABLE 1  
MEANS AND S.Ds. FOR ELEVEN VARIABLES.

(Forty-five Boys and Girls.)

Variable		Mean	S.D.
1	Speech at 2½	40.622	25.899
2	Vocabulary at 5	32.600	9.797
3	Language at 5	36.133	11.024
4	Vocabulary at 8	17.933	5.079
5	Reading at 8	30.578	20.736
6	Vocabulary at 10	38.489	14.173
7	Reading at 10	56.156	21.015
8	Written vocabulary at 10	12.733	8.365
9	Subordination index at 10	21.333	10.844
10	Composition rating at 10	46.467	21.290
11	Non-verbal intelligence at 8	21.289	4.829

TABLE 2  
INTERCORRELATIONS (PRODUCT-MOMENT) OF ELEVEN VARIABLES.  
(Forty-five Boys and Girls.)

Variables	2	3	4	5	6	7	8	9	10	11
1	.484	.458	.518	.575	.602	.594	.408	.320	.527	.206
2		.793	.752	.569	.742	.644	.365	.217	.575	.587
3			.646	.480	.600	.613	.352	.374	.638	.527
4				.645	.782	.706	.464	.240	.495	.525
5					.717	.832	.583	.330	.655	.395
6						.756	.579	.303	.624	.573
7							.503	.359	.732	.529
8								.417	.661	.249
9									.627	.209
10										.448
11										—

All  $r$ 's  $\geq .288$  significant .05 level.

All  $r$ 's  $\geq .372$  significant .01 level.

TABLE 3  
MEANS AND S.Ds. BY SEX.

Variable	Boys (N.22)		Girls (N.23)	
	Means	S.D.	Means	S.D.
1	42.455	26.454	38.870	25.233
2	34.409	9.552	30.870	9.715
3	36.000	11.123	36.261	10.927
4	19.636	4.923	16.304	4.676
5	32.955	22.618	28.304	18.475
6	42.136	14.265	35.000	13.171
7	58.636	22.896	53.783	18.738
8	12.955	6.772	12.522	9.641
9	19.091	10.808	23.478	10.438
10	44.364	21.597	48.478	20.795
11	22.682	5.226	19.957	3.983

TABLE 4  
INTERCORRELATIONS.  
(Boys, N.22.)

Variables	2	3	4	5	6	7	8	9	10	11
1	.502	.530	.581	.470	.677	.604	.691	.472	.656	.232
2		.805	.745	.575	.783	.700	.512	.216	.720	.554
3			.662	.450	.598	.614	.445	.300	.743	.565
4				.617	.774	.699	.606	.404	.564	.520
5					.679	.910	.602	.365	.680	.484
6						.754	.676	.566	.800	.683
7							.534	.426	.737	.559
8								.517	.642	.241
9									.579	.376
10										.618
11										—

All  $r$ 's  $\geq .404$  significant .05 level.

All  $r$ 's  $\geq .515$  significant .01 level.



TABLE 5  
INTERCORRELATIONS.  
(Girls, N.23.)

Variables	2	3	4	5	6	7	8	9	10	11
1	.460	.388	.466	.701	.529	.580	.222	.207	.414	.150
2		.812	.748	.552	.682	.574	.278	.309	.493	.598
3			.716	.529	.650	.629	.299	.459	.535	.547
4				.689	.756	.733	.411	.255	.555	.426
5					.767	.715	.610	.361	.669	.224
6						.761	.555	.167	.536	.353
7							.514	.357	.774	.464
8								.384	.705	.286
9									.669	.167
10										.348
11										—

All  $r$ 's  $\geq .396$  significant .05 level.

All  $r$ 's  $\geq .505$  significant .01 level

#### V.—THE WORK OF THE SAMPLE.

The work of the sample (the subjects of the longitudinal study) was analysed in various ways. It was also related to existing and new data connected with their linguistic development in other fields. (v. correlations, Tables 2, 4 and 5) and to their general adjustment, as before.

(a) *The Subordination Index.* "The intimate relationship between language development and thought processes makes a study of language development important," says LaBrant, and she finds in the subordination index a useful measure of maturity in written language and one which is somewhat parallel to sentence length as an indication of oral language development. To obtain the index the number of dependent predicates is expressed as a percentage of the total number of predicates in each child's composition. In the present sample, the number of dependent predicates ranged from 0 to 51. The mean was 21.333 ( $\sigma 10.844$ ). The girls excelled the boys, the respective means being 23.478 ( $\sigma 10.438$ ) and 19.091 ( $\sigma 10.808$ ). Correlations with other variables throw further light on the index from the developmental angle. From these it would seem to be somewhat specifically related to written language skill, in that it correlates +.627 with the children's ability in composition (as judged by impression marking) but much lower (+.320 and +.374) with oral ratings at 2½ and 5.0. However, recency and distance in time may also be factors here. The index has its lowest correlation with the Raven Matrices Test 1947 (administered when the children were 8). It gives  $\rho +.209$  overall, with boys' results showing  $\rho +.376$  and girls'  $\rho +.167$  only.

(b) *Vocabulary Studies.* The extent and quality of an individual's vocabulary provides one of the most obvious clues to his linguistic competence. Vocabulary in written composition can be studied in various ways but very long passages are normally deemed necessary (Chotlos). Here, Chotlos and Johnson's (1944) methods were applied in a modified form to the analysis of each child's two compositions.

*Core and fringe vocabulary.* The number of different words used by a writer may be taken to indicate the verbal resources at his command and the extent to which he is stimulated to make use of them. His whole vocabulary will never be displayed but it may be sampled. For this purpose, the different verbs used by the children were studied. Each separate word of verbal type was counted, in the same way as by Burroughs (1951) and Mein and O'Connor (1960) in their word counts. In the two compositions, the forty-five children used 268 different verbs. A 'core' of twenty verbs was in almost universal use. The 'fringe' verbal vocabulary (all words not in the core) varied from 0 to 20 in individual compositions. 123 verbs appeared only once. The overall average 'fringe' was 6.4. There was little difference between the sexes (boys used 6.5, girls 6.3 on average) but more between the subjects, factual subjects producing 5.8, and imaginative 6.95 fringe verbs. Scores based on the amount of fringe vocabulary correlated rather lowly (+.365 and +.464) with earlier vocabulary measures at 5 and 8, but better with vocabulary at 10 (+.579), and with composition skill (judged by impression marking) at 10, +.661. Except on this last item, boys' correlations are all higher than girls.

*Type-token analyses.* Another approach to the study of vocabulary was made by type-token analysis. The proportion of different elements in oral vocabulary has been found to reflect different levels of maturity. (McCarthy, 1930, Mein and O'Connor, 1960). For this purpose, a count was made of the different nouns, verbs and adjectives appearing in the first hundred words of each composition. Little difference was found between the proportion for nouns and verbs, or according to sex or subject, for these types. Adjectives, however, represent a more advanced and differentiated sort of vocabulary. From 0 to 9 different adjectives occurred in individual compositions and the average percentage was between 3 and 4. Boys had 1 per cent. more on factual subjects and girls the same amount more on imaginative ones. For the most part, the adjectives were of a pedestrian sort ('big', 'long' and 'little' were the outstanding favourites,) but they were not idle embellishments. There were 154 different adjectives altogether, fifty-three were exclusive to factual titles and sixty-four to imaginative titles. Colour adjectives were frequent. Childish or slang adjectives (e.g., smashing) were rare. So were highly discriminating descriptions ('fast-moving', 'weather-beaten'). Emotionally toned adjectives (e.g., 'tremendous') were rarer still. There was some evidence to suggest that the use of adjectives relates meaningfully to written language development at this stage, but longer written samples would be required to confirm this.

(c) *Impression-marking assessment.* The order of merit on the two topics as established by three expert examiners who did not know the children and who read their work in verbatim typescript, correlated +.722. In calculating the inter-correlations with other linguistic variables marks obtained on both compositions were combined. In general, the composition rating correlated reasonably well with other variables, with the exception of non-verbal intelligence (+.448). It also accorded quite well with test measures of vocabulary at 5 (+.575) and 8 (+.495) and 10 (+.624), and even better with reading at 8 (+.655) and at 10 (+.732). Oral language at 5 correlated +.638 with composition ratings (boys +.743, girls +.535). The usual tendency for more recent measures to correlate more highly than less recent is marked, but the whole body of intercorrelations seemed to justify the view that the different aspects of linguistic development are related.

(d) *Adjustment and Achievement.* In 1962, some interesting observations had been made on the relationship appearing to exist between the children's



adjustment (as measured by the Bristol Guide and by clinical assessment), and their progress in reading. To obtain further information on this relationship, Bristol Guides were again completed by the class teachers (in all cases different from those previously involved). As a result, five children who had previously obtained scores indicating maladjustment were found to be clear of this. Of the ten who now had scores showing unsettlement or maladjustment, seven were new cases. In eight of the ten cases, home conditions were extremely disturbed. Table 6 shows their scoring in respect of reading and composition :

TABLE 6  
TEN 'MALADJUSTED' CHILDREN AND THEIR ATTAINMENT SCORING.  
(Mean Maladjustment Score 19.6,  $\sigma$  3.872.)

$\sigma$ Score	Reading	Composition
+1 .....	3	1
Average .....	2	4
-1 .....	5	5

Approaching the problem from another angle, the average Bristol Guide scores of those who rank low in reading and composition was compared with the average adjustment score of the forty-five children. Table 7 shows the results. A tendency for low attainments and poor adjustment to occur together appears again, but the halo affecting teachers' judgments might, in part, explain this. Certainly not all the children would be deemed clinically maladjusted, and the average scores indicate 'unsettlement' rather than serious maladjustment as the tendency.

TABLE 7  
LOW ATTAINMENT AND SCORING ON 'MALADJUSTMENT.'

	$\sigma$ Scores		Whole Group
	Reading - 1	Composition - 1	
Mean Maladjustment Score	13.2	11.1	7.5
N .....	10	10	44
$\sigma$ .....	9.110	8.944	11.135

## VI.—CONCLUSIONS.

As a general conclusion it emerges clearly that children's skill in written composition is closely related to their other linguistic skills, past and present, but composition presents special problems to the child and develops far more laboriously than speech. As would be expected, the linguistic variables correlate more highly with each other than with non-verbal intelligence. Correlations are, in general, higher throughout for boys than for girls. If not due to sampling errors, this might indicate a difference in the verbal gifts of the two sexes which

would be worth exploring. The main results, however, may be summarised as follows :

- (1) Positive correlations were found between ratings for written composition at 10 and other linguistic variables at 2½, 5 and 8.
- (2) Correlations tended to be highest between measures currently established.
- (3) Low attainments on linguistic subjects occurred in half the cases where maladjustment was indicated by the Bristol Guide.
- (4) Disturbed home conditions were found in eight out of ten cases where maladjustment was so indicated.
- (5) Low performers on linguistic subjects had higher average scores for maladjustment than the group as a whole.

ACKNOWLEDGEMENT.—Grateful acknowledgement is made of the help of numerous parents, teachers and colleagues, and especially of the labours of the raters, Mrs. Cashdan, Dr. Coltham, and Mrs. Stevens.

#### VII.—REFERENCES.

- BILBROUGH, H. (1955). *The Effect of the Title upon Performance in English Composition*. M.Ed. Thesis (unpublished). University of Manchester.
- BURROUGHS, G. E. R. (1957). *A Study of the Vocabulary of Young Children*. Birmingham Institute of Education (Educational Monographs No. 1). Edinburgh: Oliver and Boyd.
- BUSHNELL, P. P. (1930). *An Analytic Contrast of Oral with Written English*. Teachers' Coll., Columbia University.
- CHOTLOS, J. W. (1944). Studies in Language Behaviour, IV.—A statistical and comparative analysis of written language samples. *University of Iowa Psychological Monographs*, 56, 2, 77-111.
- FORD, C. T. (1954). Development in written composition during the primary school period. *Brit. J. Educ. Psychol.*, 24, 1, 38-45.
- JOHNSON, W. (1944). Studies in language behaviour, I.—A program of research. *University of Iowa Psychological Monographs*, 56, 2, 1-15.
- MCCARTHY, D. A. (1930). *Language Development of the Pre-School Child*. University of Minnesota Press.
- NICE, M. M. (1925). Length of sentences as a criterion of a child's progress in speech. *J. Educ. Psychol.*, 16, 370.
- O'SHEA, M. V. (1907). *Linguistic Development and Education*. Macmillan.
- LABRANT, L. L. (1933). A study of certain language developments of children in grades four to twelve, inclusive. *Genet. Psychol. Monographs*, 14, 5, 387-489.
- MEIN, R., and O'CONNOR, N. (1960). A study of the oral vocabularies of severely sub-normal patients. *J. Ment. Defic. Res.*, 4, 2, 130-145.
- MORGAN, G. A. V. (1956). Verbal abilities in primary school children. *Durham Research Review*, 2, 7, 97.
- NISBET, J. D. (1955). English composition in secondary school selection. *Brit. J. Educ. Psychol.*, 25, 1, 51-4.
- PENFOLD, D. M. E. (1956). Symposium: The use of essays in selection at 11+. I.—Essay marking experiments: shorter and longer essays. *Brit. J. Educ. Psychol.*, 26, 2, 128-136.
- SAMPSON, O. C. (1956). A study of speech development in children of 18-30 months. *Brit. J. Educ. Psychol.*, 26, 194-201.
- SAMPSON, O. C. (1959). The speech and language development of 5-year-old children. *Brit. J. Educ. Psychol.*, 29, 217-222.
- SAMPSON, O. C. (1962). Reading skill at 8 in relation to speech and other factors. *Brit. J. Educ. Psychol.*, 32, 12-17.
- SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1961). *Studies in Spelling*.
- SMITH, M. E. (1927). An investigation of the development of the sentence and the extent of vocabulary in young children. *University of Iowa Stud. Child Welfare*, 3, 5, 7-92.
- WISEMAN, S. (1949). The marking of English Composition in grammar school selection. *Brit. J. Educ. Psychol.*, 19, 3, 200-209.
- WISEMAN, S. (1956). Symposium: The use of essays in selection at 11+. III.—Reliability and validity. *Brit. J. Educ. Psychol.*, 26, 3, 172-9.

(Manuscript received November 28th, 1963.)



# AWARENESS OF REVERSIBILITY: ITS EFFECT ON PERFORMANCE OF CONVERSE ARITHMETICAL OPERATIONS

BY RIVKA R. EIFERMANN AND DALIA ETZION

(*Department of Psychology, The Hebrew University, Jerusalem, Israel*)

**SUMMARY.** Why is the operation of addition easier than that of subtraction, even to adults and even when the sums involved do not exceed 10? Addition and subtraction are converse operations and it is surprising, therefore, that adults, who undoubtedly comprehend their reversible nature, find one of the operations more difficult to carry out than the other. Our assumption was that this fact resulted largely from the way the two operations were taught at school, where addition was more thoroughly trained than was subtraction. In order to test this assumption, we created a system isomorphic to the number system of from 1 to 10, a system based on the differences between various lengths of coloured rods. We then taught the system to a group of sixteen students, in a way that did not give preference to the operation of addition over that of subtraction. The result was, that when the subjects were required to carry out the two operations on the newly learnt system, the difference in response time to addition as against subtraction problems was significantly reduced when compared with the same difference obtained in solving problems with numbers. It may thus be possible that a teaching method which emphasizes the reversible nature of arithmetical operations may enable a more comprehensive and flexible approach to arithmetic.

## I.—INTRODUCTION.

Of all characteristics of thought operations 'reversibility' is the most basic for Piaget's conception of the development of intelligence. Piaget describes the properties of reversibility thus:

"Every change is reversible. Thus, the two classes of relations just combined may be separated again and, in mathematical thought, each original operation of a group implies a converse operation (subtraction for addition, division for multiplication, etc.). This reversibility is no doubt the most clearly defined characteristic of intelligence, for although motor functions and perception are capable of combination, they remain irreversible" (1947, pp. 41-42).

This paper attempts to clarify the effect of learning procedures on the ability to carry out equally well two converse operations. For our experiments we have chosen the converse operations of addition and subtraction. They are at the core of arithmetical reversibility. Yet intelligent adults seem to prefer addition over subtraction when solving problems in which either operation can be used. Moreover, addition problems are significantly easier to carry out than subtraction problems—even when the problems (or solutions) never contain a number larger than ten (Eifermann, in preparation).

Since addition and subtraction are converse operations and intelligent adults undoubtedly comprehend their reversible nature, the problem of why one should be easier than the other suggests itself. Our assumption was, that learning methods in childhood were largely responsible for the phenomenon. This assumption was based on the observation that the way most of us learn the two operations seems clearly to favour addition over subtraction. Counting

upwards is the earliest operation performed. Addition is the operation first taught, on the basis of counting upwards. Subtraction is taught at a later stage, and it is taught on the basis of addition. Subtraction is, therefore, also less repeatedly and thoroughly trained.

Hence, we hypothesized, that if we trained people in a way that would not favour the one operation over the other, we would succeed in equalizing their difficulty.

The system we presented consisted of units of length (cardboard strips) isomorphic to the number system, from 1 to 10. The way of acquainting subjects with the system was to present them with the relations of inclusion existing between the units of the system and at the same time avoiding any grading according to increasing or decreasing magnitudes. When the relationships between the lengths (differentiated through their colours—one colour for each length) were mastered, subjects were given a set of addition and a set of subtraction problems to solve, with the colour names, representing lengths, and then with numbers. Their ability at solving addition and subtraction problems (as measured by time and error scores) was determined, and so was the relative difficulties of the tasks.

Since our subjects were adults, we did not expect to eliminate the effects of transfer from the familiar number system to colour entirely. Our hypothesis, therefore, was, that *the difficulty of subtraction relative to addition would be significantly smaller with numbers than with colours* when the same subjects carry out all tasks. The measure of difficulty was response time over number of correct responses.

## II.—THE EXPERIMENT.

### (i) *Subjects and materials.*

*Coloured cardboard strips:* Ten differently coloured cardboard strips, 2 cm. wide, each longer than the preceding by 2 cm.—starting with a grey 2 cm. and ending with an orange 20 cm. strip. In this way a set was constructed, isomorphic to a number set of from 1 to 10. The strips, in increasing order of magnitude, had the following colours: grey, red, light blue (which has its specific name in Hebrew), pink, green, yellow, blue, black, brown, orange.

Colours were used for the sake of convenience only. They provided names to the differently sized strips and were thus randomly selected, having no intrinsic value as they have in Cuisenaire and Gattegno's (1952) system (see Discussion).

*Question cards:* Questions, representing all relations between the strips were printed each on a separate card. The questions were formulated as, "how many reds are there in a pink?" the answer being, "two reds." The answer to each question was printed on the opposite side of its card.

*Problems of addition and of subtraction:* There were four sets of fifty problems, each set presented on a separate sheet. These were, addition problems with numbers; the same problems with colours; subtraction problems with numbers; the same problems with colours. They were arranged as follows:

*Addition:* there are twenty pairs of reversible addition problems such as  $3+6=9$ ,  $6+3=9$ . These were divided into two columns, one of each pair in a column. The remaining five single problems (i.e.,  $5+5$ ,  $4+4$ ,  $3+3$ ,  $2+2$ ,  $1+1$ ) were presented in each of the two columns. The order of the problems within each column was also random. In this way columns of equal difficulty were obtained.



*Substraction*: Here there were also fifty problems, arranged in two columns of twenty-five. The forty paired problems were such as,  $9-5=4$ ,  $9-4=5$ . The five problems not paired were,  $10-5$ ,  $8-4$ ,  $6-3$ ,  $4-2$ ,  $2-1$ . The problems were distributed in the same random way as on the addition sheet.

The subjects were first year students of the Psychology Department of The Hebrew University of Jerusalem, five men and eleven women. These sixteen participants reached the critical test phase of the experiment. An additional three dropped out during earlier phases.

(ii) *Experimental procedure.*

Each subject had two hourly sessions with the experimenter, on successive days. The experiment consisted of three major parts:

*Practising relations.* During the first session the subject learnt relations of inclusion holding between the individual lengths of the ten different cardboard strips.

The experimenter described the purpose of the experiment to the subject thus:

"I have in my hand a set of cardboard strips varying in length and colour. Each length has its own colour and no other. That is, there are no two strips of different lengths whose colours are the same and there are no two colours whose strips are of equal length. All strips of equal length have the same colour, and *vice versa*.

"Your task is to learn and memorize the relations between the lengths of the different strips. The strips are coloured in order to help you identify them.

"The learning method will be as follows: I shall state the relationship between two strips (how many times one 'goes into' the other) and will present the strips to you for verification. Try to remember each relation presented by heart, because later on you will have to state what the relations between the strips are—without resort to the strips.

"Throughout the experiment we shall name the strips according to their colours.

"Is everything clear to you?"

The relations between the lengths of the strips were thus defined by the number of times a particular strip was longer than another, to which it was compared, e.g., "one red strip is as long as two grey strips put end to end" or, "one blue strip is as long as three red strips with an additional one grey strip." After presenting all possible relations with the strips, the experimenter provided the subject with the question cards (see Material). The subject was asked to go through all questions, answer them, and check his answers. The planned learning procedure lasted until the subject could state the relations of inclusion between the strips without seeing the strips.

At this point the first experimental session was terminated. It lasted up to one hour—in one case as little as twenty minutes. For reasons of economy, only subjects who succeeded in getting through this session in at most one hour, were invited for a second session. Three subjects did not reach this criterion and did not continue with the experiment.

*Revision and examination.* The subject went quickly over the relations learnt on the previous day, to refresh his memory. He was then presented with the forty-five question cards. But this time the experimenter presented them, one at a time, and recorded the subject's response time to each card. All subjects were presented with the cards in the same order. The criterion for continuing to the critical test phase (predetermined by the pilot study) was, that the mean response time per question should not exceed 5 seconds, and that the number of errors should not exceed five. All subjects examined reached the criterion.

*Test phase.* While the previous, practice phases, emphasized the relations of inclusion, this critical phase consisted of another task. The subject was asked to answer, in writing (up to this stage all responses were oral), four sets of fifty problems each. These were :

- (a) Addition problems with colours (e.g., "grey + red = light-blue").
- (b) Subtraction problems with colours (e.g., "red - grey = red").
- (c) Addition problems (the same as in (a)) with numbers.
- (d) Subtraction problems (the same as in (b)) with numbers.

In order to counter-balance the effect of order of subtraction and addition, the sets were presented to half the subjects in the order (a), (b), (c), (d); for the other half the order was : (b), (a), (d), (c). For each set, time taken to complete each of the two columns of twenty-five problems was recorded separately. The first column was considered practice, the second—critical. Subjects were given the opportunity to rest after each column; these rest periods lasted up to 2 minutes, according to request.

After completing the four sheets of problems, the purpose of the experiment was explained and subjects' comments were recorded (comments made during the experiments were also noted). Subjects were specifically asked not to talk about the experiment.

### III.—STATEMENT OF RESULTS.

- (i) *The measure of difficulty—response time per correct answer.*

The data obtained from each subject were : Response times to eight columns of twenty-five problems, and the corresponding number of errors. Of these, the first column on each sheet was considered the practice column and our hypothesis was tested on the four remaining critical columns.

It should be noted that degree of difficulty of a task is measurable by a combination of at least these two variables. The following measure, suggested by Halevi\* was used :

$$\frac{\text{response time (over twenty-five problems).}}{\text{number of correct responses (out of twenty-five)}}$$

(This measure represents, in fact, mean time spent on a correct response. Thus, will the score of a subject who acts quickly but whose error score is high, be equal to that of another subject who acts slowly but does not make mistakes. Of course, in so far as a person acts slowly and also makes many mistakes, his score will be high—he does not act efficiently; when he acts quickly and without error, his score will be low—he acts efficiently.)

Table 1 gives mean correct response times in seconds for each of sixteen subjects on addition and on subtraction problems with colours and with numbers.

\* We are indebted to Mr. H. Halevi, for suggesting this measure to us.



TABLE 1

MEAN RESPONSE TIME IN SECONDS PER CORRECT ANSWER ON ADDITION AND SUBTRACTION PROBLEMS WITH COLOURS AND WITH NUMBERS (N = 16).

Subject Number	Colours		Numbers	
	Addition	Subtraction	Addition	Subtraction
1	3.97	4.39	0.85	1.19
2	4.08	4.87	1.19	1.76
3	7.08	4.36*	0.63	0.75
4	4.62	4.65	0.92	1.17
5	4.10	5.09	1.00	0.93*
6	5.36	5.42	0.70	0.98
7	4.60	4.47*	0.71	0.73
8	5.73	5.02*	0.93	1.11
9	3.29	4.29	0.70	0.76
10	4.91	4.93	0.82	0.89
11	4.16	4.11*	0.71	0.91
12	4.83	4.80*	0.71	0.82
13	4.53	4.39*	0.89	0.86*
14	4.30	4.27*	0.62	0.78
15	4.16	6.05	1.00	1.17
16	4.54	4.97	0.64	0.80
Mean . . . . .	4.64	4.76	0.81	0.98

\* Subtraction easier than addition.

(ii) *Preliminary tests—differences in scores between addition and subtraction.*  
Before testing our hypothesis as stated in the Introduction, we wanted to test on the present data our finding of previous studies, that addition was significantly easier than subtraction. (a) *Numbers*: Wilcoxon's matched pairs signed-ranks test (Siegel, 1956) for related samples was calculated on the differences in scores between addition and subtraction. The results were significant ( $p < 0.01$ ). That is, subtraction was significantly more difficult than addition according to our measure. (b) *Colours*: The same test calculated on the difference between addition and subtraction with colours, was not found significant.

As can be seen in the subtraction column with colours in Table 1, the scores for subtraction are smaller than those for addition in seven (starred) out of sixteen cases. This is true only in two cases in the corresponding number column.

(iii) *Testing the hypothesis—differences in scores between addition and subtraction: colours as compared with numbers.*

Wilcoxon's matched pairs signed-ranks test for related samples was calculated on the difference between addition and subtraction on the colours as against the number scores. For each subject, each of the two differences was weighed against his mean response time.\* The calculated difference was found significant in the predicted direction ( $p < 0.01$ ). Thus, as predicted, the difference in scores between subtraction and addition was significantly greater with number problems than it was on the same problems presented with colours.

\* The weighing was necessary to enable direct comparison between two sets of scores of different orders of magnitude (see Table 1).

*Subjects' comments.*

(a) *Translation of colours into numbers*: all subjects admitted, at the end of the experiment, that they had translated to some extent. According to reports, translation always meant breaking up the particular colour into the smallest unit (grey), e.g., 'pink' equals 'four greys.' Subjects tended to translate when the answer to the question presented (examination phase) did not come automatically. Otherwise, even where there was complete implicit knowledge of the number-colour correspondences, subjects did not resort to translation. As one subject typically explained: "I know that it is possible to relate all colours to the grey, and calculate that way. It probably makes it easier to remember the colours, but it would make me go slower on the calculations. I, therefore, do it only when I am unsure of the answers."

(b) *Addition as compared with subtraction*: in most cases subjects remarked that with numbers, they found subtracting more difficult than adding. With the colour problems they were not aware of such a difference between the two operations. But they found the second column of colour problems always easier than the first. The first column thus fulfilled its function of practice.

## VI.—DISCUSSION AND CONCLUSIONS.

The results show that a relatively brief period of learning a system in a way which does not favour addition over subtraction, suffices to reduce significantly the difference in relative difficulty of these reversible operations.

Of course, the training does not eliminate the difference completely. What remains (though not statistically significant) may be due first of all to contamination resulting from partial translation of the colours into 'greys' (the smallest unit). That such translation does occur, is evident in the reported results. However, since it appears that translation occurred only where the relations between colours were not fully mastered by the subject, this factor could presumably be more fully controlled by using more stringent criteria for mastery.

Even so, one cannot exclude at this stage, the possibility of other factors operating as well. These might be products of selective emotional associations to the two operations: e.g., addition associated with 'receiving,' subtraction—with 'giving away' (such emotional associations were found operative in a related area, of response to affirmative as compared with response to negative verbal statements; Eifermann, 1961). Also, there might conceivably exist a differential 'innate capacity' to compose structures rather than to break them up.

These and other possible factors must await investigation. But, whatever their influence, the fact that learning evidently affects the relative difficulty of subtraction as compared with addition, would seem to have both theoretical and practical implications. Theoretically, it concerns the possibility (not elaborated by Piaget in his definition of operations), of enhancing the development of the capacity of children to carry out particular operations. Practically, it implies that a method of teaching arithmetic which would not favour one operation over another could facilitate greater mobility and flexibility in carrying out arithmetical operations.

In a recent discussion on reversibility, Inhelder (1962) emphasizes that only at the formal stage of thought does complete reversibility appear. By 'complete reversibility' Inhelder means, the combination of 'negation,' or 'inversion' (i.e., "the displacement in one direction [that] can be cancelled by displacements in another direction") with 'reciprocity' or 'compensation' ("the relationship between two things and cancellation of one by the other").



She emphasizes, however, that in one area complete reversibility is obtained earlier, i.e., at the stage of concrete operations: "In the concrete level, the child's thought operations involve *either* reciprocity *or* negation; and these two are combined only in the region of numbers" (1962, p. 35).

In the light of this statement, the findings, that adults prefer addition to subtraction when given the choice (see Introduction), and that they are better at solving addition problems, stand out particularly. For clearly, the comprehension of reversibility and even the capacity to carry out reversible operations when required, does not as such confer maximal "mobility and coherence upon formal thought" (Inhelder, 1962). What is required, is active awareness and readiness to manipulate where possible and desirable.

Piaget and Inhelder's investigations of the development of the *capacity* for reversible operations, confine themselves by and large to situations in which the child is *explicitly requested* to carry out an operation and its converse, or to judge an operation carried out in his presence as reversible (see, e.g., 1941, 1958).

Demonstrations of the comprehension of reversibility when specifically required to operate in ways indicating such comprehension does not imply the capacity or willingness to utilize it, entirely on one's own initiative. But, in arithmetic, as in all intelligent action, optimal mobility implies capacity and readiness to act upon the reversible nature of the given information. Thus, for example, flexibility in working with fractions or equations is only achieved when the possibilities of cancellation, or transfer from one side of an equation to the other, are fully utilized. Teaching methods which give predominance to one operation cannot encourage thinking in reversible terms. Under such circumstances of learning, the reversible nature of operations may well remain latent in relevant situations.

The teaching method with the system developed in the present investigation may, with modifications, suggest a way by which the desirable balance in mastery and facility for manipulation may be achieved. A system which corresponds on many points to the present one, has been developed by Cuisenaire and Gattegno (1952), for teaching arithmetic in primary schools. It appears to have been successfully applied in a number of countries. The method, developed by Cuisenaire, is based on the manipulation of relationships between coloured rods of different lengths, which are grouped according to colour (and number) families. From the viewpoint of our experiment, the immediate interest in the system lies in Gattegno's statement that with it, "addition and subtraction can be taught at the same time" (1952, p. 5).

The obvious next step in the present research would be a similar study, carried out directly on children. In this, experience already gained with Cuisenaire's system might well prove enlightening.

#### VII.—REFERENCES.

- CUISENAIRE, G., and GATTEGNO, C. (1954). *Numbers in Colour*. Heinemann.  
 EIFERMANN, R. (1961). Negation: a linguistic variable. *Acta Psychologica*, Vol. 18, 4, 258-273.  
 INHELDER, B. (1962). Some aspects of Piaget's genetic approach to cognition, in Kessen, E., and Kuhlman, C., *Thought in the Young Child*. Monograph of the Society for Research in Child Development, Vol. 27, No. 2, 19-34.  
 INHELDER, B., and PIAGET, J. (1958). *The Growth of Logical Thinking in Childhood and Adolescence*. Basic Books, Inc.  
 PIAGET, J. (1941). *The Child's Conception of Number*. Routledge and Kegan Paul.  
 PIAGET, J. (1947). *The Psychology of Intelligence*. Routledge and Kegan Paul.  
 SIEGEL, S. (1956). *Non Parametric Statistics for the Behavioral Sciences*. McGraw, Hill and Co.

# THE ANALYSIS OF WORD SOUNDS BY YOUNG CHILDREN

By D. J. BRUCE

*(Department of Education, University of Cambridge)*

**SUMMARY.** This study is concerned with the development of the ability to make a simple phonetic analysis of the spoken word. It describes the performance of children at different levels of mental development on the task of analysing word sounds into two components: a letter sound and a sound unit which is the residual word obtained by elision of the letter sound from the test word. The results, which refer to the mental age range 5+ to 9+, suggest that there is a gradual progression towards accurate phonetic analysis, in the course of which children:

- (a) Come to recognise that words and sounds are interrelated.
- (b) Acquire criteria of what constitutes analysis.
- (c) Achieve positional differentiation.
- (d) Overcome the cohesiveness of the word sound pattern in their experience.

These necessary features of the ability are not attained in discrete, successive steps, but are present in increasing proportions with advancing mental age. Mental age 7+ is indicated as the level at which they first become available in sufficient degree to permit some success with the task.

## I.—INTRODUCTION.

THIS study is the first in a general programme which will be concerned with psycho-linguistic development. Much useful information is already available relating to such features as the articulation and discrimination of speech sounds, growth of vocabulary, and length and structure of verbalisations (Templin, 1957). Other studies have referred to the verbal habits of children (Woodrow and Lowell, 1916; Ervin, 1961; Palermo, 1963), and their ability to appreciate the significance of grammatical structure (Werner and Kaplan, 1950; Brown and Berko, 1960). There have been enquiries into the functional role of children's language behaviour (Piaget, 1926 and 1959; McCarthy, 1930), and cognitive studies too frequent to mention make some sort of reference to the part played by language in the developmental processes involved. But it seems true to say that until the recent impact of Russian work (Liublinskaya, 1957; Luria and Yudovich, 1959; Luria, 1963), there have been few studies available which seek to relate, systematically, the factors of linguistic development to the course of intellectual progress through childhood. It is the aim of the present programme to investigate certain of those areas of psycholinguistic development which are intimately related to intellectual attainment, in particular those areas where the child is called upon to exercise an ability to operate on his own language as object.

The chosen area of this first study is that of the child's ability to make a simple phonetic analysis of the spoken word. Although the study does not deal directly with the reading process, it is this field of intellectual attainment which supplies its basic premise. The choice was made for two reasons: (a) the acceptance of the idea that, at some stage in learning to read, it is important for the child to appreciate that the sound pattern of a word is divisible into smaller units, and that these units are common to the sound patterns of other words; and (b) the belief that, although such analysis is much discussed in



considerations of reading teaching methodology, there is a need for further evidence on the way in which the process develops. There is little doubt that the process is not one which can be immediately mastered by the young child. M. D. Vernon (1960) has suggested that such analysis is probably even more difficult than the analysis of the printed word shape into its constituent letter shapes, and has stressed the strongly holistic nature that word sounds have in the young child's experience. The consensus of opinion gained from the studies of Dolch and Bloomster (1937), Schonell (1945), and Dolch (1948) is that children below the mental age of 7 show insufficient ability of this kind to benefit from systematic phonic instruction.

On the other hand, Hunter Diack, in his book *Reading and the Psychology of Perception*, has made the interesting suggestion that a kind of aural analysis of the sounds of words—albeit not a full conscious logical analysis—is implicit in a child's earliest linguistic efforts (Diack, 1960). But what succeeds this postulated implicit analysis and precedes the attainment of fluent explicit word analysis remains obscure. As Morris (1959) has said, research data on the association between phonic ability and mental maturity are rather meagre.

From this standpoint the present study seeks to contrast the ability shown by children at different levels of mental development for the task of analysing word sounds into certain constituents. Its scope is limited to reporting and discussing factual data concerning the responses of infant school children to a single task, namely, the analysis of word sounds into two components: a letter sound and a sound unit which is the residual word obtained by elision of the letter sound from the test word.

## II.—THE WORD ANALYSIS TEST.

The test, which is presented orally, consists of thirty words—twenty-six monosyllables, three disyllables, and one trisyllable. To ensure familiarity, all the test words are drawn from the first five-hundred items in G. E. R. Burroughs' vocabulary count (1957) for 5 to 6½-year-old children. The test is administered individually, and the subject is simply asked to say what word would be left if a particular letter sound were to be taken away from the test word. Both sound and position of elision are included in the instructions for each test word.

Twenty-one of the test items yield correct residual words which are also in the first five-hundred of Burroughs' count, and nine yield residual words in the second five-hundred. For ten of the test words the suggested elision is at the beginning, for ten around the middle, and for ten at the end of the word. These three groups will be referred to as 'First,' 'Middle' and 'Last' item categories. Equal numbers of first and second five-hundred residual words occur in each group. The items are randomised according to category in the actual test list. In every case, both the sound and the correct spelling of the residual word would be produced by correct elision, i.e., this type of item is avoided: B-EAR, PENN-Y, etc.

## III.—PROCEDURE.

Before attempting the word analysis task, each child is given a standard preliminary test designed to show, in a different context, his understanding of the terms and operations involved in the word analysis instructions. This preliminary test comprises: (a) Saying a word; (b) Making a sound; (c) Indicating first, middle and last child of a group in a picture; (d) Repeating the first, middle and last of three digits spoken by the experimenter (serial order is randomised for both (c) and (d)); and (e) Demonstrating which, and how many, bricks are left when others have been separated from them by the subject.

The child is then taken through a number of examples of the word analysis task, with help and explanation, before beginning the test proper.

After the test, questions related to task performance, current teaching experience, and home reading are asked, and the child's oral spelling of some of the test items is examined. The selection includes both words correctly answered and incorrectly answered in the word analysis test. The whole session is recorded on tape.

#### IV.—SUBJECTS.

The results to be described are based on a total of sixty-seven subjects drawn from three State infant schools coming under the same local authority—thirty children from School A, twenty-five from School B and twelve from School C. Chronological ages ranged from 5 : 1 to 7 : 6. Mental ages and I.Qs. for all subjects were obtained via the Stanford-Binet Intelligence Scale, Form L-M (1960).

Subjects from School A were tested first, and in subsequent testing the effort was made to select comparable, if smaller, samples through the matching of individual cases. The variables considered were sex, chronological age, mental age, I.Q., school form, and father's occupational status.

#### V.—RESULTS.

##### (i) *Word analysis scores and mental age.*

The combined results for all three schools, given in Table 1, show a regular increase in mean score with mental age.

TABLE 1

#### WORD ANALYSIS BY MENTAL AGE.

Mean Correct Scores for Schools A, B and C.

Mental Age.....	5+	6+	7+	8+	9+
Mean Score.....	0.0	1.8	8.75	16.4	26.7
Standard Deviation ..	0.0	1.6	8.8	7.2	1.95
Number of Subjects ....	4	12	16	25	10

Taking each school separately, and using Kendall's *tau* as an index of relationship, the following highly significant values are obtained for the correlation between word analysis score and mental age: School A: +0.58; School B: +0.72; School C: +0.57. Children who had a mental age below 7 showed little or no ability to perform the task. This mental age level has been mentioned from time to time in previous research literature as a necessary qualification for carrying out phonetic analysis and for making the best use of phonic teaching, and there is nothing in the present results which obviously contradicts that view. However, there is a need for caution. Most of the children below this mental age level who were tested on word-analysis were in the lower forms. Thus, they had not so much relevant teaching experience as the children from higher forms. Now the results, as shown below, also indicate that teaching experience can have an effect on the word-analysis score, and it might be that lack of relevant experience rather than actual lack of capacity is the



cause of failure. Certainly there were some children in the lower forms who had attained a mental age of 7 and yet showed a similar lack of success. But the balance of the evidence suggests a lack of capacity for the under 7 group. Failure was not universal in the lower forms for children with a mental age of 7 and over, but there were no cases of a child with a mental age under 7 in the higher forms doing any better on the test than those of this mental age group in the lower forms.

The effect that teaching experience can have on the performance of children with a mental age of 7 and above can be seen in Table 2.

TABLE 2

## WORD ANALYSIS BY MENTAL AGE.

Mean Correct Scores and Standard Deviations for Each School.

Mental Age	School A		School B		School C	
	Mean	S.D.	Mean	S.D.	Mean	S.D.
5+	0.0	0.0	0.0	0.0	—	—
6+	2.2	1.9	1.75	1.3	1.0	1.0
7+	11.9	10.7	3.25	1.6	8.8	6.9
8+	21.5	4.0	13.7	6.5	7.7	4.4
9+	27.0	1.4	25.75	2.3	28.0	1.0

Here mean word analysis scores are given by mental age for the three schools separately, and a contrast is evident at 7+ and 8+. The schools themselves contrast in their approach to the teaching of reading. School A gives a prominent place to phonic instruction from the beginning, B attempts to suit phonics to the progress of the individual child, and C does its best to avoid phonics altogether. It is interesting to note that score equivalence is apparently regained by mental age 9+. The largest scatters within schools also coincide with mental age levels 7+ and 8+, and part of their origin is apparent in Table 3 where mean word analysis scores are given by mental age and school form.

TABLE 3

## WORD ANALYSIS BY MENTAL AGE AND SCHOOL FORM.

Mean Correct Scores for Schools A, B and C.

School Form	Mental Age				
	5+	6+	7+	8+	9+
Lower	0.0	1.8	6.0	13.5	—
Middle	0.0	—	8.2	16.0	24.7
Upper	—	2.0	18.0	20.0	27.6

It can be seen that score rises not only with mental age but, from 7, with school form for a given mental age—another indication of an effect coming from experience.

(ii) *Strategies demonstrated in response to the task.*

Looking now beyond the quantitative scores to the nature of the responses, certain differences in approach to the task can be detected, and their systematic variation with mental age suggests that they may be indicative of stages in the development of the ability to make a phonetic analysis of words. The broad pattern of strategies is common to all three schools, so results have been pooled for this description.

At the youngest mental age represented in these samples, 5+, the approach to the task is simple in the extreme. The subjects can produce a word and a sound, they have been asked to operate in these terms, and they do so. But there is no interrelation of the two such as the task demands. When a word is given as the response it is generally the test word as such, without alteration. For the greater part, though, the strategy is to give a single sound or letter name. Some 85 per cent. of the errors at this mental age level take this form, and it is noteworthy that many of these responses are not a repetition of the test sound or its letter name equivalent. The instructions are acted on insofar as something is offered which differs from what the experimenter has himself said, but that something is not a word. Rarely it is a sound or letter name which is another constituent of the test word, for example, /n/ as the response to NE-S-T, /k/ to CAR-D. Incidentally, of the two schools (A and B) contributing to this mental age group the use of unambiguous letter names is restricted to school B. Most often it is a sound or letter name unrelated to the test word. In fact, the test word remains unanalysable. The child at this mental age level approaches words and sounds as alternatives and cannot manipulate their interrelations in his operations.

These strategies are still apparent at the level of 6+, but they are supplemented by new approaches. Two prototypes of genuine analysis, in which word and sound are related, now appear for the first time. In a small number of cases the interrelation takes this form: a word is produced as requested, but in place of the desired test word minus the test sound, it is a new word beginning with the test sound that is given as the response. Examples are HAY for H-ILL and KENT for MON-K-EY. The child is unable to break up the sound pattern of the test word, but he can relate the test sound to a word in this way, by offering a parallel alternative whole word which shares the characteristic of this sound—albeit not necessarily in the same position—with the test word. The second new strategy which appears at this mental age level takes the same general form, but comes closer in its result to what has been requested in the test instructions. It may be called the technique of Substitution, and it represents the largest single category of error at 6+, 37 per cent. of the 303 errors made. Actual examples will again help to clarify what is involved in the strategy: WINK as the response to PIN-K, and LAND to HA-N-D. Although unable to carry out the requested phonetic amputation on a given word, the child relates sound and word in a way which is more in accord with the general terms of reference of the instructions than is the previously described precursor of genuine analysis. Instead of merely offering a word which shares the test sound constituent, he produces, in a substitution response, a word which shares the major part of the sound pattern of the test word. However, as the examples indicate, the result is still not a true analysis of the test word. The relation between sound and word required for this analysis task is test word minus test sound equals correct response word. In substituting, the child offers a response word which is the test word with an altered letter sound (or sounds, for sometimes the sound unit involved comprises a letter pair). As before, one may



suspect that this category of error represents an offering of one whole word for another.

Substitutions constitute a sizeable proportion of the total errors. As stated, they appear in some strength at 6+, and they have a slightly greater percentage occurrence at the next mental age level, 7+ : 41.8 per cent. But the higher mental age groups show a diminishing tendency to employ this strategy. At 8+ the figure is 23.6 per cent. and at 9+ 15.2 per cent. Not only are substitutions used less frequently by the children of higher mental age, but they are made with greater regard to the indicated position of operation. Table 4 shows the percentage occurrence of substitutions by item category for children in the two mental age groups, lower (6+ and 7+) and higher (8+ and 9+).

TABLE 4

## POSITION OF SUBSTITUTION.

Percentage Occurrence by Item Category in Schools A, B and C.

Position	Item Category			Totals
	First	Middle	Last	
Beginning :				
Higher M.A. Group	80.6 (25)	20.7 (6)	8.0 (2)	33
Lower M.A. Group	86.7 (65)	59.3 (51)	90.1 (73)	189
Middle :				
Higher M.A. Group	0.0 (0)	55.2 (16)	4.0 (1)	17
Lower M.A. Group	0.0 (0)	39.5 (34)	0.0 (0)	34
End :				
Higher M.A. Group	19.4 (6)	24.1 (7)	88.0 (22)	35
Lower M.A. Group	13.3 (10)	1.2 (1)	9.9 (8)	19
Totals :				
Higher M.A. Group	31	29	25	327
Lower M.A. Group	75	86	81	

Figures in parentheses indicate number of cases.

It can be seen that the most popular position for substitution amongst the lower group is at the beginning of the test word, irrespective of whether the instructions concerned first, middle or last sound of the word. The apparent positional correspondence of Middle item category substitutions is probably spurious. Allocation of cases within the table was made on the criterion of whether the test sound was or was not still present in the response, and although the test sound is substituted for here, in no less than thirty-one of the thirty-four cases the substitution also involves an adjacent first or last sound and so could equally well be interpreted as substitution for the beginning or end of the word. And once more the beginning is favoured, in the ratio of 30 : 1.

So the overall impression gained from this substitution data is that children in the lower mental age group tend to limit their operations to the beginning of the word—the most accessible free end of the word sound envelope. The corresponding figures for the higher mental age group show a different picture. Though the distribution of First item category substitutions remains closely

similar, there is a complete reversal of the distribution for Last item category, so that it comes to share the former's high positional correspondence. However, the Middle category continues to prove difficult for these children. There is some genuine improvement in positional correspondence, though, as before, the table figure is probably over-generous, but a marked feature of the distribution is again the increased tendency to substitute for the end of the word. This 'opening up' of the end of the word to the operations of the higher mental age subjects seems also to be reflected in Table 5, which shows the distribution of percentage correct scores by item category. The differences are not large, but it is perhaps noteworthy that the 3 per cent. superiority of First item category over Last for the children with a mental age below 8 becomes a 10 per cent. superiority of Last item category over First for those with a mental age of 8 and above, while Middle lags behind for both groups.

TABLE 5

## WORD ANALYSIS BY ITEM CATEGORY.

Percentage Correct Scores in Schools A, B and C.

	Item Category		
	First	Middle	Last
M.A. below 8 (32 Subjects) .....	21.9	10.0	18.8
M.A. 8 and above (35 Subjects) .....	63.1	57.7	73.4

To conclude the account of the substitution data, a feature to which brief reference was made previously will be described in greater detail. Substitution has been interpreted as a prototype of analysis in which the child offers a response which differs from the test word through an alteration in a sound rather than through an elision. Word and sound are interrelated, but in a way that stops short of true analysis. The change is from word to word, rather than to word out of word. It is interesting that the sound unit for which substitution is made is not always a single letter sound. Moreover, the alteration of this larger unit is strongly associated with the beginning of the word.

Of the 327 substitution responses made, 229 (70 per cent.) were for a single letter sound. All but one of the remaining ninety-eight were for a letter sound pair, and their proportionate occurrence was similar in the two mental age groups. No less than eighty-three (84.7 per cent.) of these larger unit substitutions involved the beginning of the word, though the two mental age groups showed a difference in relative frequency of occurrence of the association. Percentages of the separate totals were: higher mental age group 61.3 per cent., lower mental age group 95.5 per cent. The bulk of the disparity was again accounted for by the increased tendency of the higher mental age group to substitute for the end of the word. Not only did these subjects make more substitutions at the end of the word, but they made more of their larger unit substitutions in this position. For example, where the dubious cases of Middle item category correspondence amongst the lower mental age group were distributed in the ratio of 30 : 1 in favour of the beginning of the word, the ratio for the higher mental age group was 6 : 4 in favour of the end of the word.



Reference to the implications of the larger unit will be made later in the paper after further data have been introduced.

Summarising the data so far, it can be seen that there is a gradual progress towards genuine word analysis with increase in mental age. From the primitive approach of 5+, where word and sound are not related, a stage is reached at 7+ where there is some degree of success in correct response (mean score=8.75), and the major category of error consists of responses which go some way towards meeting the demands of the task. At 8+ and 9+ this prototype, substitution, shows greater positional correspondence, and it falls from top place in the error categorisation.

Its replacement consists of mistakes made in the course of carrying out the actual process of analysis. Reference to Table 6 will show how these faulty elisions increase proportionately with mental age, from 1.9 per cent. of the errors made at 5+ to 63.6 per cent. at 9+. This table shows very clearly that the previously noted increase in correct score with mental age is paralleled by a progressive improvement in the balance of strategies employed for dealing with the task. Failure itself becomes more analytic.

TABLE 6

## ERROR ANALYSIS.

Percentage Occurrence in Schools A, B and C.

Error Category	M.A. 5+ 4 Ss. N=104*	M.A. 6+ 12 Ss. N=303*	M.A. 7+ 16 Ss. N=316*	M.A. 8+ 25 Ss. N=339	M.A. 9+ 10 Ss. N=33
(1) Don't know . . . . .	1.0	12.2	6.3	14.5	15.2
(2) Repetition of test word . . . .	9.6	21.8	1.9	4.1	3.0
(3) Single sound or letter name . .	84.6	7.6	12.7	3.8	0.0
(4) New words beginning with test sound . . . . .	0.0	5.9	1.9	4.7	0.0
(5) Substitutions (vowel sound preserved) . . . . .	0.0	37.0	41.8	23.6	15.2
(6) Faulty elision . . . . .	1.9	4.0	22.8	36.0	63.6
(7) Words preserving initial sound —other than above . . . . .	1.0	4.3	4.4	5.9	0.0
(8) Words preserving vowel sound —other than above . . . . .	0.0	4.0	4.1	2.7	0.0
(9) Verbal associations . . . . .	1.0	1.0	1.3	2.7	0.0
(10) Additions . . . . .	0.0	0.7	1.9	1.2	0.0
(11) Unclassified responses . . . .	0.0	1.6	0.9	0.9	3.0

\* The Ns. (total numbers of errors) for M.As. below 8 do not equal total possible score—total score since, in four cases of obvious difficulty the subject was not given the full test.

The nature of the faulty elisions gives some insight into what can lie between the attainment of true analysis and correct response. These mistakes are of two main kinds: elision of a sound in the test word other than the test sound, i.e., a sound in the wrong position, and elision in which the test sound plus an adjacent sound or sounds is removed, i.e., an elision which is too extensive. Examples of the first type are NAIL as the response to S-N-AIL, and MONK to MON-K-EY; of the second type, IN for S-PIN and PAR for PART-Y. Occasionally elisions made in the wrong position also involve more than one sound, e.g., STA for S-T-AND.

Reference to Table 7 shows that about half the 215 faulty elisions coming from mental age groups 7+, 8+ and 9+ fall into the too extensive category. The table also shows that such errors are much more likely to occur when the task instructions concern the free ends of the word than when the subject is directed to the middle. The Middle category items are characterised by elisions which are erroneous in position. There is no reason to believe that this difference is an artefact arising from the phonetic structure of the particular test words. It seems to be a difference in error genuinely dictated by item category.

TABLE 7

## TYPES OF FAULTY ELISION.

Percentage Occurrence by Item Category in Schools A, B and C.

Type	Item Category			Totals
	First	Middle	Last	
Wrong position . . . . .	15.6 (10)	69.8 (67)	36.4 (20)	97
Too extensive . . . . .	81.3 (52)	24.0 (23)	61.8 (34)	109
Wrong position and Too extensive . . . . .	3.1 (2)	6.2 (6)	1.8 (1)	9
Totals . . . . .	64	96	55	215

Figures in parentheses indicate number of cases.

We have again, but now associated with true analysis, the two features of the substitution data, viz. : a lack of positional correspondence in relation to Middle category items, and the treatment of more than one letter sound as a unit when the free ends of the word are involved. The occurrence of these features in the context of attempts at genuine elision warrants a firmer interpretation than would have been possible before. Lack of positional correspondence now clearly implies that there is greater difficulty experienced in dealing with the test sound when it occurs around the middle of the word. Whether this is due to difficulty in location, or to the resistance provided to elision by the interior of the word sound envelope cannot be decided on the basis of the present data. The occurrence of elisions which are too extensive indicates that even when true analysis has been achieved, there is still difficulty experienced in separating the letter sound constituents of the whole word, and the nature of their distribution implies that it is the free ends of the word in particular which encourage the taking of the larger unit. Since these children are undertaking genuine analysis, and, in the higher mental age groups at least, are dealing with parts of the word for which they show considerable positional accuracy, it may be concluded that they actually perceive these letter sounds in a unitary fashion.

Examples with a high relative frequency of occurrence are /fr/, /br/, /st/, /sp/, /n/ plus terminal consonant, and units terminating in /i/.

(iii) *Word analysis and oral spelling.*

On the completion of the Word Analysis test, each child was requested to spell out loud some of the test words. Wherever possible, these items were



selected to give equal representation of item categories and of correct and incorrect response to the test itself, and the following data are based on those cases which fulfilled the above criterion. The purpose behind this subsidiary task was, of course, to establish the relation between ability to carry out the word analysis task and ability to put the words concerned into a letter by letter form. The children were left free to employ either letter names or letter sounds in their spelling. Most of the children used the former, and there was no difference in the distribution of results for the two types of spelling. As implied by Table 8, which gives the results for School B, the association between this type of word analysis and oral spelling was surprisingly precarious.

TABLE 8  
ORAL SPELLING AND WORD ANALYSIS TEST PERFORMANCE IN SCHOOL B.

		Word Analysis		Totals
		Right	Wrong and Don't know	
Oral Spelling ..	Right (including phonetically adequate) .....	37	25	62
	Wrong and Don't Know .....	21	29	50
	Totals .....	58	54	112

Even with the allowance of phonetically adequate spelling as correct, and a one-tailed test, the degree of departure from a chance distribution was significant only at the 5 per cent. level for schools A and B, and for school C it failed to meet the level of significance. In fact, the combined results showed some 37 per cent. of correctly analysed words met with spelling failure and 44 per cent. of incorrectly answered test words were spelled correctly.

The achievement of correct oral spelling is not then a guarantee of successful analysis of a word into letter sound and residual unit. Or, putting the conclusion another way, letter by letter knowledge cannot be identified with other forms of word analysis that may be assumed to be involved in learning to read.

#### VI.—DISCUSSION AND CONCLUSIONS.

The aim of this study was to examine the ability shown by children at different levels of mental development for the task of making a simple phonetic analysis of spoken words. The results support previous findings which have suggested that mastery of such analysis takes time to develop. They indicate: (a) that a certain basic level of mental development is necessary before the child can analyse words in this way; (b) that other forms of word and sound interrelation are used before true analysis is attained; (c) that the process of actual analysis is affected by positional and holistic factors. They also indicate that teaching experience can influence the demonstrated level of analytic skill once the necessary basic mental age is attained.

Children with a mental age below 7 scored an average of 1.4 on the thirty items of the word analysis test, actual scores ranging from 0 to 4. Though the majority of these children came from the reception classes, higher forms were

represented in this mental age group and the individuals concerned shared in the low level of performance. Additional cases falling into the category of mental age below 7 and higher form, but not included in the samples for reasons concerned with matching, showed similar low scores. There was no doubt that some members of the group had had more experience of phonic instruction than others, yet the group showed a uniform absence of analytic ability, indeed of an analytic approach. In place of analysis, children at this level engaged in a variety of strategies, ranging from a predominant use of single sounds or letter names by the youngest in mental age to a predominance of substitution responses amongst the oldest. The majority of the latter responses at this mental age level take the form of a 'rhyming word', and it is perhaps worth noting that the children themselves are thus engaging in an approach which has been suggested by some investigators, e.g., Gray, Dolch, Harris, Watts, as suitable for early informal phonic training. There is no reason to suppose that the actual teaching experience received by these children necessarily forced them into this mode of response.

The association between error categories and mental age throws some light on the difficulties that are experienced in breaking down the word sound pattern, and the way in which those difficulties are gradually surmounted. The first hurdle would seem to be that of gaining the realisation that sounds and words are inter-related. As has been pointed out, the responses of children of mental age 5+ showed little appreciation of relationships between the two. When this realisation has been gained, by mental age 6+, a second difficulty is encountered in the form of relationship that is demanded by the task. Children of mental age 6+ achieve a relation between sound and word in a variety of ways other than elision. Broadly, they offer phonetic resemblance instead of phonetic analysis. Sometimes it is the test sound which is taken as the common link, sometimes, as Table 6 shows, it is another constituent of the test word—an initial or a vowel sound. Most often the response is a word which shares more extensively in the phonetic characteristics of the test word—a substitution. This response by phonetic resemblance centres almost exclusively on the beginning of the word as the determinant of operations, thus implying a third difficulty, that of position.

Children of higher mental age persist in these non-analytic approaches, but the ratio to analytic errors drops steadily, and even the non-analytic substitutions undergo a progressive improvement as approximations to the requested outcome, showing greater positional correspondence. From being predominantly rhyming responses the substitutions move on to constitute a more sophisticated group of alterations in which more attention is paid to the part of the test word concerned. In particular, the end of the word is gained as an area on which operations can be carried out. Such substitutions seem to be at the threshold of true analysis, and they are, of course, produced by children who are at the same time giving a substantial number of genuine analytic responses to other test items. But even at this stage (mental age 8+ and 9+) the problem of position continues to prove troublesome when the child is directed to the middle of the word, and elision is carried out much more readily on the free ends of the word. There would seem to be three possible sources of the greater positional difficulty of Middle category items: (a) the child may find it harder to locate the relevant sound in this position; (b) the middle of the word sound pattern may be more resistant to break-down than the beginning and the end; (c) this item category demands both analysis and synthesis in order to obtain the residual word. Any, or all, of these factors could direct the child's operation to another position.



The fourth, fundamental, difficulty which underlies the extended course of development of this ability stems from the holistic nature of word sounds in the young child's experience. Even those elisions which are carried out in the correct position often bear traces of it. Though children from a mental age of 7 on can disrupt the cohesion of a word's sound elements, the separation of these elements does not appear on an 'all or none' basis. Certain conjunctions of letter sounds are still experienced in a unitary fashion at 8+ and 9+. The reality of the difficulty is clearly expressed in the words of one subject (mental age 8:6), when faced with the task of eliding /s/ from NEST: "I can't actually do it. You see I can't say the last letter without the middle." He was successful, but the separation of /s/ from /t/ presented a considerable problem.

To summarise, this study reveals four principal barriers which must be surmounted before accurate phonetic analysis of words is possible:

- (1) Acceptance of words and sounds as exclusive categories.
- (2) Lack or confusion of criteria of what constitutes analysis.
- (3) Deficiency in positional differentiation.
- (4) The cohesiveness of the word sound pattern in the child's experience.

Mental age 7+ marks the point at which a real start is made on overcoming these difficulties, but the study suggests that even at a mental age of 9+ children are not entirely clear of the last three.

ACKNOWLEDGEMENT.—My thanks are due to the Heads of the schools concerned for their generous co-operation and help in this research study.

## VII.—REFERENCES.

- BROWN, R., and BERKO, J. (1960). Word association and the acquisition of grammar. *Child Development*, 31, 1-14.
- BURROUGHS, G. E. R. (1957). *A Study of the Vocabulary of Young Children*. Oliver and Boyd.
- DIACK, H. (1960). *Reading and the Psychology of Perception*. Peter Skinner.
- DOLCH, E. W. and BLOOMSTER, M. (1937). Phonic readiness. *Elem. School J.*, 38, 201-205.
- DOLCH, E. W. (1948). *Problems in Reading*. Champaign, Ill.: The Garrard Press.
- ERVIN, S. M. (1961). Changes with age in the verbal determinants of word association. *Amer. J. Psychol.*, 74, 361-372.
- LIUBLINSKAYA, A. A. (1957). The development of children's speech and thought. In *Psychology in the Soviet Union* (Ed. Brian Simon). Routledge and Kegan Paul.
- LURIA, A. R., and YUDOVICH, F. Ia. (1959). *Speech and the Development of Mental Processes in the Child*. London: Staples Press.
- LURIA, A. R. (1963). The role of speech in the formation of temporary connections and the regulation of behaviour in the normal and oligophrenic child. In *Educational Psychology in the U.S.S.R.* (Ed. B. and J. Simon). Routledge and Kegan Paul.
- MCCARTHY, D. (1930). The language development of the pre-school child. *Inst. Child Welfare Monogr. Ser.*, No. 4. Minneapolis: University of Minnesota Press.
- MORRIS, J. M. (1959). *Reading in the Primary School*. Newnes.
- PALERMO, D. S. (1963). Word associations and children's verbal behaviour. To appear in *Child Development and Behaviour*. Vol. 1. Academic Press.
- PIAGET, J. (1926 and 1959). *Language and Thought of the Child*. Routledge and Kegan Paul.
- SCHONELL, F. J. (1945). *The Psychology and Teaching of Reading*. Oliver and Boyd.
- TEMPLIN, M. C. (1957). *Certain Language Skills in Children*. Minneapolis: University of Minnesota Press.
- VERNON, M. D. (1960). The development of perception in children. *Educ. Res.*, 3, 2-11.
- WERNER, H., and KAPLAN, E. (1950). Development of word meaning through verbal context: an experimental study. *J. Psychol.*, 29, 251-257.
- WOODROW, H., and LOWELL, F. (1916). Children's association frequency tables. *Psychol. Monogr.*, 22, No. 97.

(Manuscript received December 13th, 1963.)

## VIII.—APPENDIX.

The words used in the Word Analysis Test are listed below in the order of presentation. The words 'middle,' 'first' and 'last' after the stimulus words indicate which part of the word is to be elided.

1. S-T-AND	<i>Middle.</i>	11. S-TOP	<i>First.</i>	21. THIN-K	<i>Last.</i>
2. J-AM	<i>First.</i>	12. FAR-M	<i>Last.</i>	22. P-LATE	<i>First.</i>
3. FAIR-Y	<i>Last.</i>	13. MON-K-EY	<i>Middle.</i>	23. S-N-AIL	<i>Middle.</i>
4. HA-N-D	<i>Middle.</i>	14. S-PIN	<i>First.</i>	24. B-RING	<i>First.</i>
5. STAR-T	<i>Last.</i>	15. FOR-K	<i>Last.</i>	25. PIN-K	<i>Last.</i>
6. NE-S-T	<i>Middle.</i>	16. C-OLD	<i>First.</i>	26. LE-F-T	<i>Middle.</i>
7. F-ROCK	<i>First.</i>	17. PART-Y	<i>Last.</i>	27. CAR-D	<i>Last.</i>
8. TEN-T	<i>Last.</i>	18. WE-N-T	<i>Middle.</i>	28. S-P-OON	<i>Middle.</i>
9. LO-S-T	<i>Middle.</i>	19. F-R-OG	<i>Middle.</i>	29. H-ILL	<i>First.</i>
10. N-ICE	<i>First.</i>	20. N-EAR	<i>First.</i>	30. EVER-Y	<i>Last.</i>



# PERSONALITY RATINGS OF ADOLESCENTS : A STUDY IN A COMPREHENSIVE SCHOOL

By H. J. HALLWORTH

*(Department of Education, University of Birmingham)*

**SUMMARY.** Ten housemasters in a comprehensive school rated all 14-year-old boys and girls within their houses on fourteen personality traits. Measures were also obtained, for the same pupils, of verbal and non-verbal intelligence, English and Arithmetic attainment, family size and ordinal position in family. A correlation matrix of all variables was computed for each sex within each house. The ten matrices for each sex were combined, a principal components analysis was obtained for each of the two combined matrices, and the axes were rotated to the Varimax and Quartimax criteria. Six factors were obtained and their interpretation is discussed.

## I.—INTRODUCTION.

WHEN teachers' ratings of their pupils' personality traits are used on school record cards it is commonly assumed that each trait is independent of every other trait. It is, of course, recognised that a 'halo' effect is possible (Thorn-dike, 1920), but this is regarded as an unfortunate artifact. The results of recent empirical studies, however, cast doubt upon these assumptions. Several questions are raised. Are we justified in regarding such trait ratings as each measuring a separate aspect of personality? Do they, in fact, give only one overall assessment of personality in the form of a 'halo effect'? Are they most appropriately regarded as organised into several groups of traits?

One of the first researches into personality ratings of children was carried out by Burt (1915). Other work up to 1953 has been summarised by Eysenck (1953), and a more recent survey has been made by Warburton (1962).

The present article offers new empirical evidence derived from ratings of the kind which teachers are frequently asked to make. The ratings were made in particularly favourable circumstances and other measures were also used to act as markers.

## II.—SUBJECTS.

The ratings were made by the ten housemasters of a comprehensive school in an industrial area of central England. The school is acknowledged to be one of the most successful of its kind in that it pays special attention to the individual child. It is organised into ten houses, each with its own separate building where mid-day meals are taken and regular house meetings held. The housemasters are specially appointed from a wide selection of applicants because of their interest in children as persons, and have every opportunity and encouragement to become well acquainted with their house members. They were, therefore, in a much better position than a class or subject teacher to assess their pupils on a range of personality traits covering behaviour in all aspects of school life. It was the headmaster's opinion that whatever groups of ratings were obtained elsewhere, the traits would in this case be used independently because of the housemasters' knowledge of the children.

Following consultation with other members of staff who knew the children well, each of the housemasters rated all children in his house who were aged 14+ years. Boys and girls were rated separately in order to avoid spurious intercorrelations of traits (Hallworth, 1961a). In total, ratings were obtained for 138 boys and 140 girls.

### III.—MEASURES.

#### (a) *Ratings.*

The personality traits rated were as follows:

- |                                 |                                |
|---------------------------------|--------------------------------|
| (1) Emotional stability.        | (7) Sociability.               |
| (2) Trustworthiness.            | (8) Self-assertion/Submission. |
| (3) Persistence.                | (9) Maturity.                  |
| (4) Co-operation with teachers. | (10) Popularity.               |
| (5) Cheerfulness.               | (11) Confidence/anxiety.       |
| (6) Sense of humour.            | (12) Spontaneity/withdrawal.   |

To these were added (13) Attainment in school work and (14) Ability in games. Ratings were made on a five-point linear scale, and for each trait a brief description was provided for the extremes of the scale. All the members of the group were rated on one trait before the housemaster proceeded to the next trait. The mean of the scale was assumed to be the mean of the group being rated, and ratings were distributed in accordance with the percentages of the normal curve. By such means it was again hoped to avoid a spurious halo effect.

#### (b) *Other Measures.*

Four measures of intelligence and attainment were available for all pupils in the investigation. These were a Raven's Matrices score (1956 revision), the test having been given by the school staff a few weeks before the ratings were obtained; and Moray House verbal intelligence, English and Arithmetic quotients, obtained as part of the secondary school entry procedure. Particulars of family size and ordinal position in family were also obtained.

### IV.—ANALYSIS OF THE DATA.

In each of the ten houses, boys and girls were considered separately. There were, therefore, twenty different groups. Following inspection of the measures for normality of distribution, a correlation matrix was computed for each of these groups. The ten matrices for each sex were then combined using *z*-scores to give the combined matrices in Table 1. A principal components analysis was obtained for each matrix, six components being extracted. These included all components with a latent root greater than unity and accounted for 81 per cent. of the variance in the boys' matrix and 79 per cent. of the variance in the girls' matrix. Finally, for each set of six components, the axes were rotated twice, once to the Varimax criterion and once to the Quartimax criterion. Since the two rotations gave almost identical results in each case, the Varimax factors only are given in Table 2.\*

\* All computations were performed on an electronic computer using programmes written by the author (see Hallworth, 1962).



TABLE I  
CORRELATION MATRICES FOR TWENTY VARIABLES.

The matrix for boys is above the diagonal, that for girls below the diagonal.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1		65	63	59	44	11	27	13	60	43	45	16	39	32	25	20	17	16	-01	-15
2	67		69	66	30	-02	28	-03	50	39	22	18	49	19	29	22	18	21	-11	-02
3	62	78		93	30	01	31	24	56	40	41	26	45	41	28	30	27	24	07	00
4	48	74	66		34	01	28	-08	46	44	19	15	27	19	15	06	05	03	15	15
5	52	32	46	41		65	68	43	25	47	60	66	13	15	09	09	07	11	13	07
6	29	21	31	31	72		73	49	10	42	52	64	06	12	07	05	06	05	-02	-08
7	30	24	33	26	63	59		55	26	68	58	72	17	33	14	08	04	05	03	00
8	-05	-10	10	05	46	61	52		33	40	67	66	28	48	25	29	27	34	15	03
9	43	43	44	34	26	26	10	24		58	41	49	37	32	34	33	30	-11	-10	
10	58	44	49	32	65	64	62	36	37		54	48	27	51	17	08	05	03	00	10
11	48	29	37	26	58	54	61	58	46	60		67	27	29	19	19	15	16	04	06
12	36	54	59	44	28	33	16	19	46	38	75		37	34	28	70	68	66	-20	-14
13	28	15	27	32	43	57	47	52	25	52	55	56		21	21	25	23	25	06	14
14	26	41	33	28	16	22	03	10	18	25	12	14	67		63	70	57	65	-14	-03
15	25	45	46	43	20	22	13	20	33	25	15	26	73	16		58	84	81	-09	-02
16	22	41	39	30	24	18	28	24	32	21	31	74	29	58	85		79	73	-16	-10
17	22	41	41	39	30	23	10	19	27	31	17	20	72	16	58	85			-07	-03
18	23	41	42	34	19	23	10	19	11	-10	-11	-13	-23	-25	-17	-19	-30	-14		58
19	-18	-08	05	-05	-14	-10	-08	-19	-11	-10	-04	-03	-22	-06	-04	-17	-22	-14		
20	-15	-03	08	-06	-02	07	01	-10	-18	-04	-05	-03	-22	-06	-04	-17	-22	-14		

## KEY TO VARIABLES.

1. Emotional stability.
2. Trustworthiness.
3. Persistence.
4. Co-operation with teachers.
5. Cheerfulness.
6. Sense of humour.
7. Sociability.
8. Self-assertion/submission.
9. Maturity.
10. Popularity.
11. Confidence/anxiety.
12. Spontaneity/withdrawal.
13. Attainment in school work.
14. Ability in games.
15. Raven's Matrices.
16. Verbal intelligence.
17. English attainment.
18. Arithmetic attainment.
19. Family size.
20. Ordinal position.

TABLE 2  
LOADINGS OF TWENTY VARIABLES ON SIX VARIMAX FACTORS FOR BOYS AND GIRLS.

Varimax Factor		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	% of variance
I	Boys ..	16	09	11	09	83	90	87	60	16	54	64	82	09	14	09	06	04	06	05	01	20.64
	Girls ..	21	05	22	20	76	83	82	79	17	66	72	87	17	68	06	09	18	09	-13	05	25.00
II	Boys ..	11	20	22	00	02	00	-01	27	27	-01	32	14	79	22	80	93	87	89	-12	-01	20.48
	Girls ..	07	33	31	26	04	17	-03	16	18	17	02	17	79	16	79	89	86	88	-16	-07	20.15
III	Boys ..	78	84	85	92	31	-10	20	-13	57	40	21	08	34	17	15	07	05	02	04	00	18.80
	Girls ..	62	83	77	83	37	10	24	-23	28	29	16	-01	27	02	05	24	17	17	05	-08	15.09
IV	Boys ..	-14	-11	08	17	11	-09	-03	11	-16	-01	08	03	-18	11	-06	00	-08	02	88	87	8.63
	Girls ..	-15	02	01	-01	-03	08	-01	-12	-09	01	-04	-03	-13	-15	01	-09	-20	-03	84	86	7.94
V	Boys ..	06	03	18	03	-16	-02	25	38	26	57	22	15	17	85	16	01	00	00	-07	14	7.65
	Girls ..	60	22	11	-15	17	10	00	-29	05	44	18	-31	15	17	34	-12	-09	-02	-17	11	5.59
VI	Boys ..	28	-16	12	-12	07	-03	-13	44	50	-11	48	17	01	16	-18	09	14	11	19	-19	5.19
	Girls ..	21	13	19	02	-04	02	-14	21	86	12	43	03	24	13	-03	07	01	07	09	-15	6.11

NOTE.— For key to variables, see Table 1.



## V.—DISCUSSION OF RESULTS.

The Varimax factors obtained in the two analyses, particularly the first three factors, are remarkably alike both with respect to the variables which have high loadings and the total amount of variance extracted. They indicate that in general the housemasters made ratings in a consistent manner for both sexes.

(a) *Factor I : Sociability or Extraversion.*

For boys' and girls' analyses the first factor accounts for one-fifth and one-quarter of the total variance respectively. In both cases it has high loadings for the ratings on (6) Sense of humour, (7) Sociability, (5) Cheerfulness, (12) Spontaneity, (8) Self-assertion, (11) Confidence and (10) Popularity. This is obviously a teacher's picture of the cheerful, sociable child. Loadings for all the other ratings except one are positive but, with one exception, small. This exception is the high loading for (14) Ability in games in the girls' analysis. It appears that when differentiating between girls with regard to sociability, the housemasters took games ability into account much more than in the case of boys.

(b) *Factor II : Intelligence and Attainment.*

The second factor accounts for approximately one-fifth of all the variance of each of the two analyses, and may be identified as 'intelligence and attainment.' The variables having high loadings are (16) Intelligence, (18) Arithmetic, (17) English, (15) Raven's matrices, and (13) Teachers' rating of attainment in school subjects. Two other ratings have small but consistent positive loadings on this factor, namely, (2) Trustworthiness and (3) Persistence. Such a result could well be expected. Similarly, it is not surprising that large family size and late ordinal position have very small negative loadings: although obviously insignificant in the statistical sense, they are precisely the loadings expected in view of the finding of the 1947 Scottish survey that such measures have a general negative correlation with intelligence (Scottish Council for Research in Education, 1958).

Certain other ratings have low positive loadings in the boys' analysis. These are (11) Confidence, (8) Self-assertion and (9) Maturity. The rather low positive loading on confidence, here rated as the opposite of anxiety, is again quite consistent with results obtained by investigations of a quite different type, in which grammar school children have been shown to have a significantly lower level of questionnaire anxiety than secondary modern school children (e.g., Hallworth, 1961b).

A graphical plotting of the first two factors indicates, for both boys and girls, two well-defined clusters of variables as described above. If the axes were rotated to pass through the centre of each cluster, the two factors would be positively correlated. In other words, sociability or extraversion would have a small correlation with intelligence and attainment.

(c) *Factor III : Reliability and Conscientiousness or Emotional Stability.*

The third factor accounts for approximately 18.8 per cent. of the variance in the boys' analysis and 15.1 per cent. in the girls' analysis. High loadings are on the ratings of (4) Co-operation with teachers, (2) Trustworthiness, (3) Persistence, and (1) Emotional stability. The factor may be named either 'reliability and conscientiousness' or 'emotional stability.'

In the boys' analysis there is a somewhat lower loading on (9) Maturity. In view of earlier findings that grammar school teachers place the trait of maturity on this factor, whilst secondary modern school teachers place it, for



boys, on the sociability or extraversion factor (Hallworth, 1961a), the suggestion is that housemasters made their ratings more in the manner of the grammar than of the secondary modern school teachers. This again is perfectly consistent with results obtained in other studies, in which comprehensive schools have been shown to be more comparable to grammar than to secondary modern schools in respect of the attitudes they produce (Miller, 1961).

Again, graphical plottings of Factors I and III, and of Factors II and III, indicate that there would be in every case a slight correlation between the axes if these were made to pass through the centres of the clusters of variables. In other words, sociability or extraversion correlates slightly with reliability and conscientiousness or emotional stability; and this in turn has a low correlation with intelligence and attainment.

(d) *Factor IV: Family size and Ordinal position.*

The fourth factor has high loadings on the measures of (19) Family size and (20) Ordinal position in family. It is of little interest except to indicate the low loadings of all other variables on this factor.

(e) *Factor V: Popularity.*

The fifth factor is somewhat different in the two analyses. For the boys' analysis it extracts 7.65 per cent. of the variance and has high loadings on the ratings of (14) Ability in games and (10) Popularity, and a moderate loading on (8) Self-assertion. For the girls it extracts only 5.59 per cent. of the variance and has a high loading on (1) Emotional stability, a moderate loading on (10) Popularity and (15) Raven's Matrices, and negative loadings on (8) Self-assertion and (12) Spontaneity.

The factor suggests that the house masters regarded popularity differently in the two sexes: for boys, they associated it with games ability, and for girls, with emotional stability and a somewhat quiet personality. The rating on popularity has already been noted to have high loadings on Factor I. There are, apparently, two kinds of popular child: the sociable child, and the boy or girl described by the present Factor V.

(f) *Factor VI: Maturity and Confidence.*

The last factor extracts only 5.19 per cent. of the variance in the boys' analysis and 6.11 per cent. in the girls' analysis. However, it has fairly consistent loadings, which are in each case highest for the ratings of (9) Maturity and (11) Confidence. There are also consistent positive loadings for (8) Self-assertion and (1) Emotional stability. Apparently, this can best be described as a factor of social maturity.

In view of the accumulating evidence that mean scores on self-rating anxiety scales are lower for higher age levels, it is of interest to note that the housemasters associated maturity with confidence.

## VI.—CONCLUSION.

The first conclusion from this investigation is that when the housemasters rated their pupils on the fourteen traits, they did not use these traits independently. Despite the fact that they were assumed to have a detailed knowledge of their pupils, their ratings were largely grouped into two factors tentatively labelled 'sociability' or 'social extraversion,' and 'emotional stability' or 'reliability and conscientiousness.' Further, this occurred whether they were rating boys or girls. The two factors were in both cases correlated with each other, and with the factor of intelligence and attainment. The implication is that there was some degree of halo effect in the ratings, the pupil with highest



intelligence and attainment being attributed with other desirable personality traits.

An earlier and more limited analysis of teachers' ratings yielded the same factors of extraversion and reliability and conscientiousness. The traits of popularity and confidence fell between these factors and it was predicted that they might form the basis of further factors. The trait of maturity fluctuated in position (Hallworth, 1961a). In the present research the inclusion of measures of intelligence and attainment, the extraction of all latent roots greater than unity, and rotation to an approximation of simple structure, gave three further factors, numbered II, V and VI. Factor II may be identified as 'general intelligence' and has low loadings on all ratings except attainment in school subjects. Factor V is one of the predicted factors, identified by the rating of popularity; Factor VI is the other, based upon the ratings of maturity and confidence.

It is apparent that Factors I and III have much in common with the dimensions of introversion-extraversion and neuroticism-emotional stability described by Eysenck (1953) and with the comparable second order factors obtained by Cattell (1957). It may also be remarked, however, that these two factors have much in common with the two main dimensions of meaning obtained by Osgood (1957). When raters judge people on a variety of seven-point scales identified by opposing pairs of adjectives, it is found that two main dimensions are obtained, one of 'activity-potency' or 'dynamism' and one of 'evaluation.' It could be argued that, when the housemasters rated their pupils on the traits used in this investigation, they applied two dimensions of meaning within the context of school life and, in effect, asked themselves first, "How dynamic is this pupil in the social life of the school?" i.e., how humorous, sociable, cheerful, spontaneous and self-assertive, etc., is he? and second, "How do I evaluate him as a pupil?" i.e., how co-operative, trustworthy, persistent and stable is he?

The factors may indicate aspects of children's personality. Alternatively, they may indicate the manner in which teachers organise their perceptions of the children, or the way in which they organise the meaning of trait names. The two kinds of interpretation are not necessarily inconsistent.

#### VII.—REFERENCES.

- THORNDIKE, E. L. (1920). A constant error in psychological ratings. *J. Appl. Psychol.*, 4, 25-29.
- BURT, C. (1915). The general and specific factors underlying the primary emotions. *Brit. Assoc. Ann. Rep.*, 84, 6945.
- EYSENCK, H. J. (1953). *The Structure of Human Personality*. London: Methuen.
- HALLWORTH, H. J. (1961a). Teachers' personality ratings of high school pupils. *J. Educ. Psychol.*, 52, 297-302.
- HALLWORTH, H. J. (1961b). Anxiety in secondary modern and grammar school children. *Brit. J. Educ. Psychol.*, 31, 281-291.
- HALLWORTH, H. J. (1962). Electronic computer programmes for psychologists. *Bull. Brit. Psychol. Soc.*, 48, 16-21.
- MILLER, T. W. G. (1961). Values in the comprehensive school. *University of Birmingham Inst. of Educ. Monograph No. 5*.
- SCOTTISH COUNCIL FOR RESEARCH IN EDUCATION (1958). *Educational and Other Aspects of the 1947 Scottish Mental Survey*. U.L.P.
- CATTELL, R. B. (1957). *Personality and Motivation Structure and Measurement*. London: Harrap.
- OSGOOD, C. E., SUCI, G. J., and TANNENBAUM, P. H. (1957). *The Measurement of Meaning*. Urbana: Univ. of Illinois Press.
- WARBURTON, F. W. (1961). The measurement of personality. *Educ. Res.*, 4, 2-17.

(Manuscript received 19th December, 1963.)

# A COMPARISON OF TWO APPROACHES TO THE TEACHING OF SPELLING

By P. S. FREYBERG

(*Massey University of Manawatu, New Zealand*)

**SUMMARY.** The spelling performance of a large group of 9-11-year olds who had been taught from teacher-provided lists, was measured both by a dictated-word test and from a sample of written expression. Two years later, the same tests were administered to a comparable group from the same schools, all except two of these having by that time adopted an alphabetic spelling list from which children compiled their own learning lists. The results showed that the better spellers achieved more according to both measures through compiling their own lists, but that the poorer spellers made fewer errors on the dictated-word test when using teacher-provided lists. It is suggested that the two approaches might well be combined in the case of duller or younger children.

## I.—INTRODUCTION.

DURING the two decades from 1940 to 1960, spelling was taught in New Zealand primary schools almost exclusively by the list method: that is to say, the teacher assigned each day or week a group of words to be learnt by all pupils in his class, the words being taken from a published list such as that of Schonell (1932). In the latter part of this period, however, teachers became increasingly dissatisfied with the published lists which, having been prepared under conditions obtaining more than twenty years earlier in another country, appeared no longer to be entirely suitable for use in their schools. Accordingly, it was suggested that the New Zealand Council for Educational Research should compile a list which would better meet the spelling needs of children in this country.

In view of both the criticism that teacher-provided lists cannot cater for the needs of individual children, and the research evidence that a considerable amount of spelling is learnt incidentally, the N.Z.C.E.R. decided to present its selected words in such a way that children would build up their own individual spelling lists from the words they used in their everyday writing. Thus, the words finally chosen (N.Z.C.E.R., 1960) were presented in alphabetic order, to enable a child to find those he required quickly and easily. To prevent the child's learning task from becoming excessive, the list was divided into seven levels, according to the relative frequency with which the words were used, the level being shown by a number placed next to each word in the alphabetic list.

Briefly, the procedure advocated in the teachers' manual (Arvidson, 1960) is that the child should consult the alphabetic list whenever he is not sure of the spelling of a word he wants to use in his written work, or whenever the teacher has indicated a word he has spelt wrongly. If the indicated level of the word is at or below that at which he is currently working—which has previously been decided by the teacher on the basis of special tests sampling the words at each level—then the child adds that word to his learning list. If the word is at a higher level, the child simply writes it correctly but makes no special effort to learn it.

Even before the alphabetic lists were published it became obvious that they would be widely adopted in New Zealand schools, and that some effort should be made to assess any changes in spelling attainment which might occur. Adequate comparisons of the effectiveness of different teaching methods are notoriously difficult to make, because of the number of variables involved.



In this case, however, two radically different approaches to the teaching of spelling were involved: in the past, children had been provided with a set list of words by their teachers, and in future, they would compile their own individual lists. The methods by which children would *learn* the words would vary from one classroom to the next under both sets of conditions, but the introduction of the new approach, provided a valuable opportunity for a broad comparison of the effectiveness of teacher-provided lists as against individually-compiled lists.

It was, therefore, decided to measure the spelling achievement of a comparatively large group of children at a single class level, immediately before the introduction of the new spelling books and to compare this with the achievement of a similar group of children two years later. Two methods of assessment were used: the number of spelling errors on a test of forty words which were included in both the Schonell and the N.Z.C.E.R. alphabetic lists, and the number of separate words wrongly spelt during a fifteen-minute 'composition' period. Thus, the investigation was aimed at studying changes in spelling ability as revealed in everyday writing activities, as well as on a dictated list of isolated words.

## II.—PREVIOUS RESEARCH.

Criticisms of spelling programmes based on teacher-provided lists have been made periodically over the past sixty years, the first with some research backing being probably that of Rice (1897). His article was followed by a number of studies in which conflicting opinions were expressed on the value of formal teaching contrasted with the incidental learning of spelling. Attention then became focussed on the words contained in spelling lists, rather than on the value of the lists themselves. An accessible account of the early work in this field has been prepared by Patterson (1961).

While the use of scientifically-selected lists was obviously an improvement on previous methods of choosing the words which children should learn, a number of writers still maintained that the procedure of requiring all children in a class or group to learn the same words was educationally unsound. Investigations such as those of Curtis and Dolch (1939) and Kyte (1948) confirmed the amount of incidental learning which takes place before children are presented with the words in a formal spelling lesson. Objections to the waste of time involved have been at least partially met by the adoption of pre-testing to find out those words which children should concentrate on learning. But, although this method has been demonstrated to be superior to the study-test method for most children, the latter may be better suited to the needs of young children—under 9 years of age—and duller children (Patterson, 1961, p. 80).

The importance of incidental learning has led to efforts to devise ways of taking full advantage of it. Hildreth (1955) devised a list of words arranged alphabetically with each word assigned to one of six learning levels, according to frequency of usage by elementary school children. Her work provided the basis for an experimental scheme in schools of the City of New York, and later for the N.Z.C.E.R. alphabetic list (1960). Freyberg (1957) produced an alphabetic list for junior (primary) school children without indicating levels, as he considered that if a word was used frequently enough to justify its inclusion in the basic list, then it should be learnt by children as soon as they required it in their everyday writing.

There has also been some evidence, however, that the learning of words as they are required for everyday writing is not in itself sufficient. Thus, Tyler (1939) concluded that some systematic teaching of spelling was necessary,



particularly to establish efficient learning habits. Deacon (1956) compared the results obtained from testing 414 second and third grade children (8 and 9-year-olds). He found that those who had been taught with teacher-provided lists showed significantly greater gains in reading tests, spelling tests and in accurate spelling in story writing, but that those who had learnt their words from individually-compiled lists developed sentence ability in their writing to a greater measurable degree.

### III.—PRESENT INVESTIGATION—METHODS.

#### (a) *Subjects.*

All children in Standard 3 classes in State primary schools in a small provincial city, and in a neighbouring farming town, took part in the study. The testing was conducted by the head teachers of the fourteen schools concerned during the first week of November (i.e., towards the end of the school year in the Southern hemisphere), when the average age of the children was approximately 10 years.

On the first occasion, when all the schools had been using teacher-provided lists throughout the year, 838 children were tested. Two years later, results were obtained from 864 Standard 3 children in the same schools, but at this time, two of the schools had not yet changed over to use of the alphabetic list. The results for these two schools were, therefore, analysed separately from the others on both occasions.

The number of children provided by the twelve schools which made the change was sufficiently large for each of the testings (760 and 771, respectively) to justify the assumption that no significant differences in the overall level of intelligence, or in the effectiveness of the teachers concerned, would have been likely to occur within two years. Any differences revealed by the results should thus reflect the effects of changing from one approach to the teaching of spelling to another. In the case of the two schools which retained teacher-compiled lists, however, the numbers were small (seventy-eight and ninety-three, respectively) and there is evidence that in one of these schools, the average level of intelligence happened to be considerably lower on the second occasion than on the first. Unfortunately, this fact blocked an additional avenue of comparison between the two approaches.

#### (b) *Tests and procedures.*

(i) For the dictated-word test, forty words which appeared in both the Schonell and the N.Z.C.E.R. alphabetic lists were selected. Both lists give the level at which the average Standard 3 pupil will be working, and the words were chosen to sample this level as well as those above and below it for each list. As the level to which a word is assigned varied from one list to the other, there were nine possible categories.

The first group consisted of six words which should have been learnt by average children in previous years (i.e., in Standards 1 and 2), according to both Schonell and N.Z.C.E.R. lists; the second, of four words occurring at Standard 3 level, for the average child, in the N.Z.C.E.R. list, but earlier in Schonell; and the third, of four words which the average Standard 3 child would not be expected to note down for learning according to the N.Z.C.E.R. list, but which he would have been taught in Standard 1 or 2, using Schonell; and so on. The nine groups thus formed the pattern shown in Table 1.



TABLE 1

TEST WORDS ACCORDING TO CLASS LEVEL.  
(Number of words in each group in parenthesis.)

		Schonell list class level.		
		S1-2	S3	S4+
N.Z.C.E.R. list class level . .	S1-2	I (6)	IV (4)	VII (4)
	S3	II (4)	V (4)	VIII (4)
	S4+	III (4)	VI (4)	IX (6)

The sampling of words in nine categories enabled the setting up of several hypotheses regarding the results which might be expected following the change from teacher-provided to pupil-compiled lists. These were:

- (1) That the proportion of errors in Group I words, which should have been learnt in previous years, according to both lists, would not vary much whichever list was used. If anything, there would be fewer errors when the alphabetic list was used, as it would provide better opportunities for revision through the looking up and noting of errors.
- (2) That the proportion of errors in words in Groups II, IV and V (i.e., earlier or current learning in both lists) would not change significantly.
- (3) That the proportion of errors in words in Groups III and VI would be expected to increase with the use of the N.Z.C.E.R. list, as the words are beyond the normal range of Standard 3 children according to that list. The words would, however, be subject to more incidental learning (through being looked up but not noted in the child's personal list).
- (4) That the words in Groups VII and VIII would be subject to fewer errors when the N.Z.C.E.R. list was used, as this list puts them at or below the level for average Standard 3 children.
- (5) That there would be little change in the proportion of errors in Group IX (outside the range of average Standard 3 children on both lists), but perhaps one could expect a slight decrease in errors due to better incidental learning with an alphabetical list.

(ii) In addition to the dictated word test, a sample of the children's written expression was obtained. The children were told to imagine that they had received an unexpected sum of a hundred pounds and were asked to write about all the things they might do with it. They were not warned of any time limit and were encouraged to continue writing about the different ways in which they might spend the money, but, in fact, they were told to stop work after exactly fifteen minutes.

It was then possible to calculate not only the mean number of spelling errors on each occasion, but also the average output of words and so the proportion of errors. No instructions were given concerning the use of dictionaries

(or of the alphabetic list on the second occasion) but teachers were asked not to help any child by spelling a word for him. Thus, this test permitted some assessment to be made of the effectiveness of the new approach towards achieving its aim of helping children avoid spelling errors in their ordinary written work.

In calculating the errors made by each child, it was decided to count only the separate words wrongly spelt. For obvious reasons, any mistakes in personal names, proprietary names and other proper nouns were excluded from the error count, but all other words were included regardless of whether or not they appeared in either of the spelling lists.

#### IV.—RESULTS.

The overall results obtained from the two administrations of the dictated-word test are shown in Table 2. It can be seen that the average number of errors increased significantly both in the case of those schools which had changed from the teacher-provided to the alphabetic list (from which children compiled their individual list) and also where no change in approach had taken place. The latter result, which was not expected, is perhaps explained by the change in average level of intelligence in the smaller group which was referred to earlier.

TABLE 2  
ERRORS ON DICTATED-WORD TEST.

	Approach to spelling	No. of children	Mean errors	S.D.	Difference in means	S.E.d	C.R.	P
Lists changed	Teacher-provided lists .....	760	12.81	9.04	+ 1.47	.50	2.93	< .01
	Individually-compiled lists .....	771	14.28	10.51				
Lists not changed	Teacher-provided lists .....	78	11.40	7.46	+ 3.88	1.33	2.92	< .01
	Teacher-provided lists	93	15.28	9.92				

In the case of those schools which had changed their approach to the teaching of spelling it will be noted that there was a substantial increase in standard deviation as well as in the mean number of errors. This fact suggested the possibility of a disproportionate increase in errors amongst the poorer spellers where teacher-provided lists had not been used. This hypothesis was borne out when the results for the upper and lower halves of the groups (according to overall test results) were re-analysed separately (Table 3).

An examination of the score distributions of the two groups of children reveals that the main source of difference between them is with the poorest spellers in each case. Where individually-compiled lists had been used, 19 per cent. of the children made twenty-five or more errors in the dictated-word test, compared with 12 per cent. for those who had been taught from teacher-provided lists. The proportion of children making thirty-three or more errors increased from approximately  $4\frac{1}{2}$  per cent. to 9 per cent.



TABLE 3

ERRORS OF TOP AND BOTTOM GROUPS WHERE TEACHING APPROACH ALTERED.

	Approach to Spelling	No. of children	Mean errors	S.D.	Difference in means	S.E.d	C.R.	P.
Upper half	Teacher-provided lists .....	380	5.78	3.10	+0.07	.23	0.30	N.S.
	Individually-compiled lists ....	386	5.85	3.27				
Lower half	Teacher-provided lists .....	380	19.94	7.32	+2.74	.57	4.85	<.001
	Individually-compiled lists ....	385	22.68	8.29				

When the nine groups of words which comprised the dictated-word test are considered separately, it can be seen that there was a significant increase in the errors made in the first six groups, but not in the three groups of words which would not yet have been taught at the Standard 3 level with teacher-provided lists (Table 4). In fact, there was a small but significant decrease in Group VIII errors (words occurring at the Standard 3 level in the alphabetic list, but at a higher level in the teacher-provided list). As hypothesized, there was also a decrease in the proportion of errors with Group VII words, but this did not reach an acceptable level of significance.

TABLE 4

ERRORS ANALYSED BY GROUPS.

Group	Errors as percentage of words tested		Difference
	Teacher-provided lists	Individually-compiled lists	
I	18.18	21.10	+2.92†
II	15.65	19.91	+4.30†
III	18.07	24.45	+6.38†
IV	18.86	24.51	+5.65†
V	15.10	22.67	+7.57†
VI	17.55	27.63	+10.08†
VII	54.64	54.31	-0.33
VIII	61.96	59.56	-2.40*
IX	59.25	60.96	+1.71

\*— $P < .05$  (one-tailed).

†— $P < .001$ .

The results obtained from the second method of assessing the changes which had taken place following the introduction of the new approach are given in Table 5. Again, these have been analysed for the upper and lower halves of each distribution, which brings out the fact that, in their everyday writing, children using the alphabetic lists produced fewer errors throughout the whole range. In those schools where the alphabetic list was not used, however, the mean number of errors increased slightly but non-significantly.

TABLE 5  
ERRORS IN WRITTEN WORK.

	Approach to Spelling	No. of children	Mean errors	S.D.	Difference in means	S.E.d	C.R.	P
Lists changed upper half	Teacher-provided lists .....	380	1.00	.78	-0.35	.05	7.00	< .001
	Individually-compiled lists ....	384	0.65	.62				
Lists changed lower half	Teacher-provided lists .....	379	5.46	3.66	-0.85	.25	3.39	< .001
	Individually-compiled lists ....	384	4.61	3.25				
Lists changed whole group	Teacher-provided lists .....	760	3.21	3.13	-0.60	.16	3.82	< .001
	Individually-compiled lists ....	771	2.61	3.03				
Lists not changed	Teacher-provided lists .....	78	2.32	2.83	+0.80	.44	1.84	N.S.
	Teacher-provided lists .....	93	3.12	2.83				

The decrease in errors achieved by children using the alphabetic list appears to have been accomplished at the expense of the total number of words written within the fifteen-minute time limit—means, 79.87 and 70.94 for the two groups, the difference being significant at the .001 level. However, the proportion of errors to words written also decreased, and this result too was statistically significant. In the case of the schools, which did not change lists, no comparable significant differences were observed, either in the total words written or in the proportion of errors.

#### V.—DISCUSSION AND CONCLUSIONS.

The results of this investigation suggest that an approach to the teaching of spelling through having children compile their own individual lists of words to learn leads to a slight increase in the number of errors made on a dictated-word test. However, when the number or the proportion of errors made in ordinary



written work is considered, this approach may well be more efficient than that of providing a set list of words for children to learn each week.

While it is difficult to obtain an objective estimate of the amount of time spent by children on the learning of spelling under the two sets of conditions, teachers generally considered that less time was needed where children were using the alphabetic lists. Few classes maintained the daily twenty-minute study period that was an integral part of the teacher-list programme. On the other hand, children's attention was better directed towards spelling during the whole of the day's activities.

The gain in accuracy of spelling revealed in the sample of written work, can with some justification be considered more important in the long run than the comparatively small increase in errors on the dictated-word test. Consciousness of the value of correct spelling and the habit of checking words about which one is unsure are likely to lead to continued learning when formal spelling lessons, have been outgrown. If this can be accomplished, together with less expenditure of time, as appears to be the case, then the use of individual lists compiled from an alphabetic list is amply supported.

Nevertheless, in view of the previous research on this topic, one aspect of the results warrants particular attention—the tendency for poor spellers to make even more errors on a dictated-word test when using individual lists instead of teacher-compiled lists. Gates (1931) concluded that while brighter children learnt most efficiently by a pretest-study method, duller children—especially those in the younger age-groups—achieved more when they all studied a set list of words. This was, in general, confirmed by Blanchard in a study cited by Fitzgerald (1951), which also recorded a finding that individual lists led to better results in the case of brighter children in particular. The results reported above are consistent with these conclusions.

An explanation for this situation can probably be found in the way in which children compile their individual lists. The brighter child has more to write about, uses more words and thus checks himself (or is checked) on more words. The duller child, who in any case is more likely to have difficulty in retaining correct spellings, writes less and so gets less practice. This is true also of the young child. The danger in using the individual approach with these children is that they may consciously or unconsciously restrict themselves to a meagre writing vocabulary; this point in itself would bear further investigation as the number of errors occurring in written work would, in such circumstances, be a misleading measure of spelling ability. Both kinds of test used in this investigation may well be necessary.

There appears to be a case, then, for the continued use of teacher-provided lists for poorer spellers and perhaps for younger children. There is no reason why the two approaches could not be combined for these groups, as suggested by Deacon (1956) and the present writer (Freyberg, 1960). As far as average and better spellers of the 10-year-old age group are concerned, however, the results of this investigation support the contention that individually-compiled spelling lists are at least as efficient as teacher-provided lists, and have additional pedagogical advantages.

ACKNOWLEDGEMENTS.—The author acknowledges with gratitude the assistance given by teachers and headmasters of Palmerston North and Fielding, in conducting this investigation, and also the help of students of Palmerston North Teachers' College in marking and collating of the results.

## VI.—REFERENCES.

- ARVIDSON, G. L. (1960). *Learning to Spell*. Wellington: New Zealand Council for Educational Research.
- CURTIS, H. A., and DOLCH, E. W. (1939). Do spelling textbooks teach spelling? *Elem. School J.*, 39, 584-592.
- DEACON, L. (1956). The teaching of spelling can become too individualized. *Education*, 76, 300-302.
- FITZGERALD, J. A. (1951). *The Teaching of Spelling*. Milwaukee: Bruce.
- FLEMING, C. M. (1952). *Research and the Basic Curriculum*. London: University of London Press.
- FREYBERG, P. S. (1957). *Junior Spelling Dictionary*. London: Macmillan.
- FREYBERG, P. S. (1960). *Teaching Spelling to Juniors*. London: Macmillan.
- GATES, A. I. (1931). An experimental comparison in the study-test and test-study methods in spelling. *J. Educ. Psychol.*, 22, 1-19.
- HILDRETH, G. (1955). *Teaching Spelling*. New York: Holt.
- KYTE, G. C. (1948). When spelling has been mastered in the elementary school. *J. Educ. Res.*, 41, 47-53.
- New York City Board of Education (1951). *A Manual to Guide Experimentation with Spelling Lists A, B and C*. New York: Division of Curriculum Development.
- NISBETT, S. D. (1939). Non-dictated spelling tests. *Brit. J. Educ. Psychol.*, 9, 29-44.
- NEW ZEALAND COUNCIL FOR EDUCATIONAL RESEARCH (1960). *Alphabetical Spelling List*. Books 1 and 2. Wellington: N.Z.C.E.R.
- PATTERSON, A. C. (1961). A review of research in spelling, in *Studies in Spelling*, Scottish Council for Research in Education, Publication No. 40. London: University of London Press.
- RICE, J. M. (1897). The futility of the spelling grind. *Forum*, 23.
- SCHONELL, F. J. (1932, 1944). *Essentials in Teaching and Testing Spelling*. London: Macmillan.
- TYLER, L. K. (1939). *Spelling as a Secondary Learning*. Contrib. to Education No. 781. New York: Teachers' College, Columbia University.

(Manuscript received 18th December, 1963.)



# THE RELATIONSHIPS BETWEEN INTROVERSION-EXTRAVERSION, NEUROTICISM AND PERFORMANCE IN SCHOOL EXAMINATIONS

BY DENNIS CHILD  
(*City of Leeds Training College*)

**SUMMARY.** The Junior Maudsley Personality Inventory was given to samples of children from an urban comprehensive school (138) and a boys' public school (40). In addition, the General Anxiety Scale for children and a reminiscence test were administered in the comprehensive school. Both I.Q. measures and terminal examination results were taken from the school record cards. A teacher-rating of Extraversion-Introversion was obtained from the public school.

The results indicated that Extraversion, neuroticism and I.Q. were statistically unrelated and that there was a positive correlation between stable Introversion and attainment in school examinations. High or even moderate neuroticism or general anxiety were not advantageously related to attainment and the Yerkes-Dodson law was not detected. Reminiscence scores were negatively related to attainment which is in the predicted direction. An analysis of the mean I.Q. and Extraversion scores for promoted and demoted children showed that the former were significantly introvert whilst demoted children tended to be extravert. The public school sample showed some unusual personality trends and some teacher-ratings at this school seemed reasonably sound.

## I.—INTRODUCTION.

It is now recognised that I.Q. and attainment measures give only a limited idea of how a child might succeed in secondary school or university work. Vernon (1961) shows that 14 per cent. of children are incorrectly allocated at the 11+ as judged by their subsequent success, also, one-third of children entering a grammar school have difficulty with the work. One possible explanation of these anomalies is Alexander's X (industriousness) factor which includes influences such as home background, teaching method and personality traits. To date, none of the components of the X factor has been isolated with sufficient accuracy to be of practical use.

Several investigations have been made into the relationships between introversion, neuroticism and academic performance in adults. Some of this work is based on a theory of personality posited by H. J. Eysenck (1947, 1953, 1957) which proposes that there are dimensions of personality almost independent of each other, i.e., extraversion-introversion (E and I), neuroticism-stable (N) psychoticism and intelligence (I.Q.). In addition, he (1957) has put forward postulates about the extravert and introvert which can be interpreted to bear directly on their possible output in scholastic work. Briefly, the postulates state that (a) human beings differ in the speed and strength with which excitation and inhibition are produced and the speed with which inhibition is dissipated, and (b) individuals in whom excitatory potential is generated slowly and is relatively weak and in whom reactive inhibition is developed quickly and strongly, but dissipated slowly are thereby predisposed to extravert patterns of behaviour. Introverts follow the exact opposite of this pattern. The effect that these qualities might have on examination performance is apparent when one considers the concentration required in revision and during the examination



itself particularly in the absence of personal contact. Many objective tests have been used by Eysenck (1957) to support the postulates and these are explained fully in the reference. Some of these tests could be modified for use with children. Reminiscence is one such test and occurs after a period of learning and rest when the amount recalled is usually more than immediately after the learning period. Differences between scores have been explained in terms of reactive inhibition. As the task proceeds, inhibition builds up to varying degrees and a suitable rest period is required for complete dissipation. The greater the inhibition dissipated, the greater will be the improvement after rest. Consequently, extravert reminiscence scores should be higher than introvert scores.

Evidence is accumulating to show that neuroticism and autonomic drive are related (Furneaux, 1961). The effectiveness of this drive and the level at which it is deleterious to performance in classroom situations is far from understood. If N may be taken as a measure of drive level, there should be a similar association between the former and performance since performance, whether scholastic or otherwise, is dynamically dependent on the strength of drive. The material collected to date is not convincing. Furneaux (1956) finds a positive relation between N and attainment for university students, also (1962) that neurotic-introverts and extraverts produce better examination results than stable-introverts and extraverts. Lynn (1959) shows that N is a factor in educational success. On the other hand, several investigators (Sarnoff, 1959; McCandless and Casteneda, 1956; and Biggs, 1962) show negative correlations between anxiety and attainment. Since N and anxiety are highly correlated, it is assumed that similar qualities as psychological definitives are being measured. Eysenck has also found some relationship between anxiety and the E-I dimension.

The Yerkes-Dodson law (1908) indicates that performance in tasks improves with increase of drive up to an optimum that is characteristic of the task and its difficulty. The optimum decreases with increase of task difficulty and beyond the peak, drive becomes maladaptive. Furneaux (1961) suggests that if drive and N are related, one would expect the law to be obeyed when using measures of N as a variable with attainment and for the effect to be curvilinear. One outcome of curvilinearity is that normal correlation techniques will give lower values of  $r$  and will mask any tendencies present. Lynn and Gordon (1961), using the Maudsley Personality Inventory and Raven's Progressive Matrices, plotted N against matrices scores and obtained a curvilinear relationship. Savage (1962) also found a U-shaped relationship using similar variables.

From the foregoing, it might be suspected that in addition to I.Q. and specific abilities the variables E and N have some influence on attainment in school children. As yet, few studies have used children as subjects with these dimensions in mind primarily because no satisfactory measure was available before the introduction of the Junior Maudsley Personality Inventory (J.M.P.I.). The authors, W. D. Furneaux and H. B. Gibson (1961), constructed a questionnaire with 22 E items and the same number of N items.

In the light of the hypotheses mentioned, this study set out to investigate (a) the correlation between E, N and I.Q.; (b) the relationship between E, N and performance in school examinations; (c) the position of anxiety with respect to E and N; (d) the Yerkes-Dodson effect in school examinations; (e) the use of a reminiscence test as an alternative to introspective measures of the E-I dimension; (f) differences in the E and N scores for promoted and demoted children; (g) the comparison between public school and day school samples; (h) teacher-ratings of the E-I dimension.



## II.—THE INVESTIGATION.

(i) *Subjects and materials.*

Two studies were made. The first used 138 pupils from an urban comprehensive school. The S's were selected at random from each form in the first four years, excluding the backward forms. The selection thus represented six from each form (three of each sex) from the first eight streams and covered the age range 11 years to 15 years.

The following tests were administered: (a) the J.M.P.I.; (b) the General Anxiety Scale for Children (G.A.S.C.), by Sarason, *et al.* (1960); and (c) a reminiscence test. The latter was repeated three weeks later. Other information included the 11+ Moray House I.Q. scores and the results of the Christmas terminal examination.

The second study involved forty fourth form public school boys between 14 years and 15 years, to whom the J.M.P.I. was given.

(ii) *Experimental procedure.*

The experimenter was not known to the S's as the authors of the J.M.P.I. suggested that familiarity between tester and testee could influence the results adversely. The usual standard testing instructions were followed.

The G.A.S.C. was constructed for use with young children in America, but has been suitably adapted for English S's. Questions were read to the S's and they responded by placing a ring round the words 'yes' or 'no.'

For a reminiscence test, S's were provided with a sheet containing ten rows of thirty-six squares. They were asked to write the letters of the alphabet upside-down from right to left as quickly and accurately as possible. After thirty seconds, they were asked to continue with the next letter on the next line and so on down the page to line nine, thirty seconds being allowed for each line. A rest of ten minutes was given after line nine and the task then repeated once more on line ten. S's were not to see previous work and a cover sheet was supplied with which to hide preceding lines.

This was not an easy task to explain to children and as a safeguard the first two lines were trials thus enabling the experimenter to go round checking work. In addition, the experiment was repeated at a later date without a trial run and the scores for this run used in the analysis. The average of the last four pre-rest trials was subtracted from the score in the post-rest trial and the result taken as the reminiscence score.

Three teachers in the public school were asked to assess the S's for E by giving them a grade between 0 and 5, inclusive. This represented the scale from extreme I to E. Several teachers were asked to define E and the three chosen expressed ideas approximating to a reasonable definition. Christmas examination results were also recorded.

(iii) *Methods of analysis.*

The first group of analyses dealt with intercorrelations between N, E and I.Q. If the Eysenckian dimensions were justified, no correlation between the variables would be anticipated.

In searching for E-N-attainment relationships, it was decided to abandon correlation coefficients since these assume linearity. A procedure adopted by Furneaux (1961 and 1962) was used. Every S had a particular N and E score above or below the mean for each scale. If the two scales were arranged at right-angles to each other with the origin at the means, four quadrants would be generated and these were referred to as 'zones'. A mean attainment score was

then calculated from individual S scores within each zone. Six zones were also produced by dividing the N scale into three sections. The mid-point of the central part represented the N score for which attainment was highest, i.e., the Yerkes-Dodson optimum. The two points on the N scale required to divide it into three parts were chosen at equal and convenient distances on either side of the mid-point and gave three sections labelled HN (high neuroticism), MN (medium neuroticism) and LN (low neuroticism). The E scale was divided at the mean whence scores above the mean were considered extravert. Predictions arising from the zones depend upon the combination of factors within the zone. The groups obtained were MN.I (neurotic-introvert), MN.E (neurotic-extravert), HN.I (highly neurotic-introvert), HN.E (highly neurotic-extravert), LN.I (stable introvert) and LN.E (stable extravert).

Attainment (At.) measures were obtained by using the terminal examination marks in English, Mathematics, History and Geography. Mean scores were calculated for each of these subjects by classes and S's credited with one point for a subject score higher than the class mean. This method was thought to be comparable with the pass-fail methods adopted by some investigators with university students. Since the preparing for and the performing of scholastic examinations are, by and large, similar processes for the school subjects selected, it was considered reasonable to take the total number of passes as a measure of success in school examinations.

### III.—RESULTS.

TABLE 1

CORRELATION COEFFICIENTS BETWEEN EXTRAVERSION-NEUROTICISM, EXTRAVERSION-INTELLIGENCE AND NEUROTICISM-INTELLIGENCE FROM TWO SIMILAR EXPERIMENTS.

	E-N	E-I.Q.	N-I.Q.
Child .....	-0.07	+0.10	-0.14
Callard and Goodfellow (1962) ....	-0.10	+0.01	-0.19

Table 1 gives the intercorrelations between E, N and I.Q. for 138 S's in the comprehensive school. Similar correlations quoted by Callard and Goodfellow (1962) are appended. None of the correlates is significantly different from zero except for Callard and Goodfellow's N-I.Q. co-efficient. The sign trend is the same in the two sets illustrated.

There are no significant differences between the J.M.P.I. quoted means and the comprehensive school sample (Table 2). The E scale mean for the public school sample is significantly different from the quoted means at the 1 per cent. level and the N score is the highest met with so far.

In the zone analysis, the E scale is divided at 12.34. On plotting the At. scores against N scores, a slight maximum is reached at N=6, although it is not convincingly higher than adjacent scores. The central MN section is taken from N=4 to N=8, inclusive. The At. means are calculated for each zone as shown



TABLE 2

J.M.P.I. DISTRIBUTION DATA: A COMPARISON WITH QUOTED NORMS.

		Extraversion	Neuroticism
Furneaux and Gibson (1961) .....	Mean .....	12.39	7.35
	S.D. ....	3.46	3.54
	N .....	156	156
Child (Comprehensive School) ....	Mean .....	12.34 m=12.60 f=12.0	6.65 m=6.75 f=6.25
	S.D. ....	3.53	3.28
	N .....	138	138
Child (Public School) .....	Mean .....	10.13	7.93
	S.D. ....	2.54	3.43
	N .....	40	40

TABLE 3

ZONE ANALYSIS OF HIGH, MEDIUM AND LOW NEUROTIC-EXTRAVERSION AND HIGH, MEDIUM AND LOW NEUROTIC-INTROVERSION WITH CORRESPONDING ATTAINMENT MEANS

	HN.E.	MN.E.	LN.E.
Attainment Mean.....	1.22	1.86	2.23
N .....	23	22	22
	HN.I.	MN.I.	LN.I.
Attainment Mean.....	2.26	2.34	2.84
N .....	26	44	6

in Table 3 (maximum At. score=4). Tests of significance of difference between the At. means show that the HN.E zone mean is lower than any other mean (2 per cent. level, compared with the nearest mean in the MN.E zone). The MN.E—MN.I attainment means are also significantly different (5 per cent. level).

With the four-fold analysis (Table 4) the zones are demarcated at E=12.34 and N=6.65. Again, the mean At. score for the NE. zone is significantly lower than any other zone mean (1 per cent. level, compared with the nearest mean in the SE zone).

The anxiety scale (G.A.S.C.) correlates with the N scale to give  $r=+0.47$  and shows  $r=-0.16$  with the E scale which is significant at the 5 per cent. level. The anxiety scale was introduced to determine whether anxiety is a more efficient indicator of drive than N as suggested by some investigators. The analysis in Table 5 is similar to that of Table 3.

TABLE 4

ZONE ANALYSIS OF NEUROTIC-EXTRAVERT, STABLE-EXTRAVERT, NEUROTIC-INTROVERT AND STABLE-INTROVERT WITH CORRESPONDING ATTAINMENT MEANS.

	NE.	SE.
Attainment Mean ..	1.39	2.08
N .....	31	36
	NI.	SI.
Attainment Mean ..	2.34	2.37
N .....	38	38

TABLE 5

ZONE ANALYSIS OF HIGH, MEDIUM AND LOW ANXIOUS EXTRAVERTS AND INTROVERTS WITH ATTAINMENT MEANS.

	H.A.E.	M.A.E.	L.A.E.
Attainment Mean.....	1.29	1.59	2.41
N .....	17	28	22
	H.A.I.	M.A.I.	L.A.I.
Attainment Mean.....	2.16	2.36	2.61
N .....	19	39	18

Improvement in At. scores extends from HA to LA across the Table and from E to I down the Table. The overlap of At. scores in the zones LA.E, HA.I and MA.I might be caused by the inter-relationship between anxiety and introversion. However, low anxiety and introversion are still significantly more advantageous than high anxiety and extraversion (difference between the means significant at the 1 per cent. level).

The correlation between reminiscence scores and At. scores is  $-0.20$  and is significant at the 5 per cent. level. Since E is not suspected of having a curvilinear relationship with At., a similar analysis yields  $r = -0.18$  which is also significant at the 5 per cent. level. The direct correlation between reminiscence and E is  $+0.17$  (5 per cent. level).

TABLE 6

MEAN EXTRAVERSION, NEUROTICISM AND ANXIETY SCORES FOR PROMOTED AND DEMOTED CHILDREN.

	E	N	Anxiety
Promoted (N=17) .....	10.70	6.47	12.8
Demoted (N=14) .....	12.94	6.70	13.9
Sample Norms. ....	12.34	6.65	12.8



Two additional analyses are included from the data. In the first experiment, some children in the sample had been moved from one class to another as a result of their performance in school examinations. Table 6 is compiled of promoted and demoted S's with their corresponding E, N and anxiety scores. The means for these variables are compared with the sample norms. First forms are not considered since little movement had occurred at the time of the experiment.

Promoted children are significantly introvert (5 per cent. level) and display average neuroticism and anxiety. Demoted children tend to be extravert and more anxious than average. Unfortunately, the sample is rather small, but it might suggest another line of investigation. 77 per cent. of promoted children have I.Q. scores which are *less* than the class average *from* which they were transferred, whilst 72 per cent. of demoted children have I.Qs. *greater* than the mean for the class *from* which they were moved. Only three children were moved back at a subsequent date.

The teacher-ratings for the public school give a joint correlation with the E-I scale of +0.40 (2 per cent. level). One teacher had a correlation of +0.55.

#### IV.—DISCUSSION AND CONCLUSIONS.

The evidence presented shows that there are at least three statistically unrelated factors, viz., extraversion, neuroticism and I.Q. which, by definition, govern the way in which an individual is likely to perform a task. Eysenck's theory is thus fundamental to learning and if the effect of the personality variables could be established, the teacher would be aware not only of the differences in ability but also in the way a child, by virtue of his or her personality, might only be capable of a limited or variable output under certain conditions.

This study, along with several others, shows a positive relationship between introversion and attainment. The reasoning is that during revision, the extravert child will have difficulty in maintaining interest in what is truly a boring task. In the examination, the process of remembering and committing to paper over a fair period of time is monotonous and an inhibition-producing situation, although it may not be fatiguing in the physiological sense. No evidence has yet been provided to show whether the E score of an individual varies greatly over the period of years spent in the secondary school. If E is largely incommutable, then due emphasis must be placed upon the way in which an individual's personality might inevitably affect his performance. There is a temptation in a classroom of comparably able children to assume that each is capable of producing work of similar quality and comments on some terminal reports are reminiscent of this influence.

The analysis of promotions and demotions shows that some movements between classes, after initial placements based on the 11+, are dependent upon introversion. This in no way casts doubts on the validity of the I.Q. measures since these are not intended to forecast examination performance, although a positive correlation exists between I.Q. and attainment in examinations ( $r = +0.4$  to  $+0.6$ , Vernon, 1961), but shows that additional information such as I scores might be used to show more nearly the style of education from which the child would most benefit. The important point is that some children of mediocre I.Q. can attain well if introvert.



One possible criticism of the analysis of promotions is that the I.Q. measures were based on 11+ results and that these children are the late developers. But, the number of children promoted constitutes 14 per cent. of the whole sample which is much higher than the number normally met with in a school population. Nevertheless, it must be admitted that a proportion of these are late developers which does not, in any way, explain why their E scores are lower than the rest of the sample. Selection for the sixth form and university is worth consideration. At the moment, most sixth formers select themselves by virtue of their 'O' level results, and yet the more exacting work of 'A' level is beyond some who are so selected. Furneaux (1961a) signifies that 20 per cent. of sixth formers suffer from temperamental disadvantages leading to high failure rate in the 'A' level examination, also that a number of secondary modern school children (6 per cent.) would probably pass at least five subjects at 'O' level, given the chance. No mention is made of the traits of these children, but one wonders if the bulk of the 6 per cent. would not be introverts with mediocre I.Q. Furneaux (1962) also shows that some 60 per cent. of stable extraverts fail university examinations.

In the zone analysis, the mean At. scores improve from low to high E and from high to low N. The N-At. score trends of Table 3 do not support the idea that high or even moderate N in school children as assessed by the J.M.P.I. is more advantageous than low N. Several MN ranges from the sample mean were examined along with other arbitrary origins on the N scale, but no other tendency was discernable. The zones show that the neurotic-extravert is least able to cope with examinations whilst the stable-introvert does the best. Intermediate categories give less reliable figures and no conclusions can be drawn from them, except that the means from E to I are significantly different.

The correlations between anxiety, N and E support Eysenck's observation that general anxiety is related to both E and N. No evidence for the Yerkes-Dodson law is found in the N range used. One possibility is that examinations constitute a task of great difficulty as defined by the law. This makes the Yerkes-Dodson optimum very low, thus placing the scores in this experiment on the right hand descending limb of the curve of N against task performance. However, this interpretation means that most children in school examinations experience a high degree of maladaptive drive which is not unreasonable, considering that examinations are a task of great complexity. Another solution may be that N and anxiety as measured by introspective methods with children are not indicative of the drive level in the manner expected.

One exception to the above findings merits discussion. The mean N score for the public school sample proved to be higher than the test norms. Entrants at 13 years constitute 70 per cent. of the fourth form and, in addition to holding the 11+, they are required to pass an entrance examination consisting of papers in a number of school subjects, i.e., Mathematics, English, a language, Science, etc. Admission is, therefore, coupled with an ability to succeed in examinations. Any sample taken from the school population should yield trends in keeping with the first experiment and this certainly applies to the E scale where pupils are significantly introvert. An explanation for the high N mean is not easily found, but since it is a biased sample, there may well be an explanation.

The reminiscence test was an attempt to compare objective and introspective measures of E. The conclusion is that the test, as applied in this study, gives equally significant predictions of attainment as the E test. It was thought



that the reminiscence test might give better predictions than the J.M.P.I. However, there is the reservation that refinement of the test and its administration might give more fruitful results.

The teacher-ratings show that some teachers are fair judges of the E-I dimension provided that they understand the criteria by which the dimension is assessed.

ACKNOWLEDGEMENTS.—The author wishes to thank the publishers, J. Wiley and Sons, inc., New York, for permission to use the General Anxiety Scale for Children devised by S. B. Sarason *et al*, and W. D. Furneaux for supplying the Junior Maudsley Personality Inventory. Appreciation is also expressed to the headmasters, staff and pupils of the two schools for their willing co-operation in the experiment.

#### V.—REFERENCES.

- ASTINGTON, E. (1960). Personality assessments and academic performance in a boys secondary grammar school. *Brit. J. Educ. Psychol.*, 30, 225-236.
- BENDIG, A. W. (1960). Extraversion, neuroticism and student achievement in introductory psychology. *J. Educ. Res.*, 53, 263-267.
- BIGGS, J. B. (1962). *The Psychological Relationships Between Cognitive and Affective Factors in Arithmetic Performance*. Ph.D. Thesis, Univ. of London.
- BROADBENT, D. E. (1958). *Perception and Communication*. London: Pergamon Press.
- CALLARD, M. P., and GOODFELLOW, C. I. (1962). Neuroticism and extraversion in school-children as measured by the junior Maudsley personality inventory. *Brit. J. Educ. Psychol.*, 32, 241-250.
- CASTENEDA, A., McCANDLESS, B. R., and PALERMO, D. S. (1956). The children's form of the manifest anxiety scale. *Child Dev.*, 27, 3, 317-323.
- EYSENCK H. J. (1947). *Dimensions of Personality*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1953). *The Structure of Human Personality*. London: Methuen.
- EYSENCK, H. J. (1954). *Uses and Abuses of Psychology*. Penguin Books (Pelican series).
- EYSENCK, H. J. (1957). *The Dynamics of Hysteria and Anxiety*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1959). Comments on a test of the personality satiation-inhibition theory. *Psychol. Repts.*, 5, 592.
- FRANKS, C. M. (1957). Personality factors and rate of conditioning. *Brit. J. Psychol.*, 48, 119-126.
- FURNEAUX, W. D. (1957). *Report to Imperial College of Science and Technology*.
- FURNEAUX, W. D. (1961). Neuroticism, extraversion, drive and suggestibility. *Int. J. Clin. Exp. Hypnosis*, 9, (4), 195-213.
- FURNEAUX, W. D. (1961a). *The Chosen Few*. Oxford Univ. Press.
- FURNEAUX, W. D. (1962). The psychologist and the university. *Universities Quarterly*, 17, 1, 33-47.
- FURNEAUX, W. D., and GIBSON, R. B. (1961). A children's personality inventory designed to measure neuroticism and extraversion. *Brit. J. Educ. Psychol.*, 31, 204-207.
- HALLWORTH, H. J. (1961). Anxiety in secondary modern and grammar school children. *Brit. J. Educ. Psychol.*, 31, 281-291.
- KIENTZLE, M. J. (1949). Ability patterns under distributed practice. *J. Exp. Psychol.*, 39, 4, 532-537.
- KIMBLE, G. A. (1949). Various articles on reminiscence in the April, June and August editions of the *J. Exp. Psychol.*
- LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.*, 29, 213-216.
- LYNN, R. (1960). Extraversion, reminiscence and satiation effects. *Brit. J. Psychol.*, 51, 318-321.
- LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.
- LYNN, R. (1962). Reply to a critical note by J. R. Biggs regarding Lynn's paper in the 1961 edition of the *Brit. J. Educ. Psychol.* *Brit. J. Educ. Psychol.*, 32, 188-199.

- McCANDLESS, B. R., and CASTENEDA, A. (1956). Anxiety in children, school achievement and intelligence. *Child Dev.*, 27, 379-382.
- SARASON, S. B., *et al.* (1960). *Anxiety in Elementary School Children: A Report of Research*. New York: John Wiley and Sons.
- SARNOFF, I. (1959). Test anxiety and the 11+ examination. *Brit. J. Educ. Psychol.*, 29, 9-16.
- SAVAGE R. D. (1962). Personality factors and academic performance. *Brit. J. Educ. Psychol.*, 32, 251-253.
- VERNON, P. E. (1961). *Intelligence and Attainment Tests*. Univ. of London Press.
- YERKES, R. M., and DODSON, J. D. (1908). The relation of strength of stimulus to rapidity of habit formation. *J. Comp. Neurol. and Psychol.*, 18, 459-482.

(Manuscript received 19th December, 1963.)



## RESEARCH NOTES

### PERSONALITY AND THE MEASUREMENT OF INTELLIGENCE

H. J. EYSENCK AND P. O. WHITE\*

(*Institute of Psychiatry, University of London.*)

**SUMMARY.** A re-analysis is presented of some data purporting to show that stable children differ from labile ones with respect to the structure of their intellectual abilities. The hypothesis is supported, and additional data are presented tending to show that theories of linear independence between cognitive and non-cognitive areas may have to be supplemented by theories stressing non-linear dependence.

It is usually maintained that intelligence is statistically independent of temperamental factors such as neuroticism and extraversion and the evidence does, indeed, show little cause to doubt lack of correlation between the cognitive and the conative-affective sides of personality. (Cf. Cattell, 1963, for a recent study and discussion). However, it would be unwise to equate *statistical independence* with *lack of interaction*; most studies reported in the literature have used statistical methods based on product-moment correlations, thus setting orthogonality equal to linear independence, and failing to allow for the possibility of curvilinear regression. This failure to take into account more complex modes of causation may have arisen from the fact that psychometric procedures have been developed very much in isolation, and without connection with the large body of experimental psychology. Eysenck (1957) has argued that the study of temperament and of intelligence can be enriched tremendously by regarding the performance of personality and intelligence tests from the point of view of experimental psychology, considering it as subject to the well known laws of learning theory, and making predictions from these. The usefulness of this approach to the study of personality variables, such as neuroticism and extraversion, has been demonstrated in several publications (Eysenck, 1960, 1964). In this paper, we shall be concerned with a consideration of a similar approach to intelligence test problem solution.

It has been argued (Eysenck, 1957) that the performance of a typical intelligence test may be regarded as an instance of massed practice, in which very similar tasks are attempted repeatedly without the interposition of a programmed rest pause. Under these conditions, we would expect reactive inhibition to build up and interfere with the proper execution of the tasks. We would also expect that extraverts, liable as they are to greater accumulation of inhibition, would show work curves different from those produced by introverted subjects, an expectation shown to be verified by two experimental studies at a high level of significance (Eysenck, 1957, pp. 132-133). In another study, Eysenck (1959) predicted that "in the process of solving the sixty problems of the Morrisby Compound Series Test . . . extraverts would show greater reactive inhibition, and consequently falling off in performance during the last quarter of the test as compared with the first three-quarters." The results showed "that extraverts show greater work decrement . . . by taking longer to obtain correct solutions toward the end of the test, as compared with introverts, and by giving up more easily toward the end" (p. 592). (At the beginning of the work, extraverts were significantly quicker than introverts.)

As regards neuroticism, it has become customary to regard this variable as in some ways being synonymous with drive; this supposition, taken together with the Yerkes-Dodson Law, may be taken to imply the likelihood of a curvilinear relationship between intelligence and neuroticism, extremely high and extremely low values of *N* being equally incompatible with high scores on intelligence tests. Lynn and Gordon (1961) have reviewed some of the literature on this point, and they have also

\* Supported by the Medical Research Council.



reported an experiment of their own which strikingly (and significantly) supported this prediction, although on only a rather small number of subjects. Their findings on extraversion were indeterminate, probably because they purposely used a very short version of the Matrices test, thus making it impossible for any large amount of inhibition to accumulate. Furneaux (1962) has also shown in connection with the prediction of success of university students that simple linear correlations are much less informative than hypothesis-directed investigations into personality-intelligence relations of a more complex character. We may conclude from this brief review that there is ample evidence to suggest that temperamental and cognitive aspects of personality may not be as unrelated as has often been supposed, and that specific hypotheses about their interrelations can be formulated on the basis of modern learning theory and its extension to personality.

One such extension of the traditional approach may be made in the field of factor analytic determination of personality structure. The problem which arises may perhaps be put as follows: When a factor analysis is carried out of personality inventory scales, a number of factors, such as extraversion, neuroticism, etc., usually results (Eysenck, 1960); similarly, when a factor analysis is carried out of intelligence test scales, a number of factors such as verbal ability, perceptual ability etc., usually result (Vernon, 1958). These factors are independent, in the first case of intelligence, in the second case of neuroticism or extraversion, as long as we preserve the rule that we are only concerned with linear relations. But we may enquire whether similar factors and relations would emerge if we extracted personality factors from populations differing in intelligence level, or intelligence factors from populations differing in degree of neuroticism, say.

A recent study by Shure and Rogers (1963) has attempted to answer the first question. They administered the eighteen scales of the California Psychological Inventory (CPI) to three student groups differing without overlap in I.Q. level, and then intercorrelated and factor analysed the resulting scores for the three groups separately. They found that while there was considerable overall similarity in the solution, the total factor variance associated with their neuroticism factor dropped by over 30 per cent. in going from the high ability group to the low ability group. (The sum of squared loadings is, respectively, 5.18, 4.64 and 3.48 for the three groups.) No such change was observed in their extraversion factor, the sum of squared loadings being 3.46, 3.76 and 3.17, respectively, for the three groups. While confirmation would, of course, be essential before too much credence can be given to this finding, it would appear that factorial studies of personality may not give invariant results under change of ability level.

The other problem raised is perhaps even more important from the educational point of view; would factorial studies of abilities be invariant under change of personality composition of the groups under analysis? It is with this question that this paper is particularly concerned.

The only paper concerned specifically with this problem is one recently published by Lienert (1963). His work is based on 1,003 school children with a mean age of between 15 and 16; three-fifths of the children were male. These children were administered thirteen intelligence tests of the Thurstone (1938) type, constituting the so-called Leistungsprüfungsystem of Horn (1962a). Also administered was a personality questionnaire modelled after Eysenck's (1953) M.M.Q. by Horn (1962b) which gives a measure of neuroticism and also contains a lie scale. Seventy-seven subjects were excluded from the analysis because they had not completed all the tests or because of unusual lie scale scores. Of the remaining subjects, 259 labile and 262 stable children were selected as constituting the 25 per cent. highest scoring and lowest scoring subjects, respectively, on the neuroticism scale. There were no differences between the groups in age but there were more girls in the labile group. However, Lienert was able to show in a preliminary factor analysis that sex had no effect on the factorial structure of the tests. A product moment correlation of the summed standard scores on the thirteen tests with neuroticism gave a value of  $-0.16$ ; while statistically significant because of the large numbers this is for practical purposes equivalent to a finding of orthogonality between the two variables



Separate matrices of intercorrelations were calculated for the labile and stable subjects, respectively, and split-half reliabilities were calculated for all the tests for the two groups. Reliabilities did not differ, but the average intercorrelation of the tests was slightly and significantly higher for the stable group (.33 as opposed to .27).

TABLE 1

Test :	Stable					Labile			
	I	II	III	h <sup>2</sup>	Lien- ert	I	II	h <sup>2</sup>	Lien- ert
1. Discovery of rules (reasoning) . . . . .	.62	.05	.18	.63	.67	.25	.58	.47	.54
2. Problems (reasoning) . .	.63	.07	.13	.63	.84	.46	.40	.49	.60
3. Word knowledge (verbal comprehension) . .	.03	.81	.08	.37	.66	.88	-.09	.45	.77
4. Word completion (verbal compr. and closure)	.49	.42	-.03	.57	.69	.27	.13	.22	.14
5. Word fluency (verbal compr. and fluency)	.07	.84	.04	.40	.90	.82	-.02	.45	.78
6. Rotation (spatial orientation) . . . . .	.18	.06	.43	.39	.45	-.01	.45	.25	.24
7. Brick-counting (spatial orientation) . . . . .	.07	.00	.76	.43	.63	.07	.71	.44	.53
8. Plane counting (spatial orientation) . . . . .	.19	.01	.65	.48	.65	-.02	.53	.29	.32
9. Hidden figures (spatial orientation and clos.)	.56	.06	.21	.60	.69	.32	.38	.39	.36
10. Hidden pictures (closure) . . . . .	-.03	.27	.40	.27	.62	.04	.47	.29	.30
11. Words (word fluency) . .	.24	.38	.02	.36	.56	.56	.05	.34	.42
12. Word beginnings (word fluency and verbal comprehension) . . . .	.51	.25	.05	.56	.56	.62	.15	.44	.48
13. Counting (number) . . .	.22	.18	.17	.34	.40	.31	.02	.19	.21
	1.00	.33	.42			1.00	.32		
	.33	1.00	.17			.32	1.00		
	.42	.17	1.00						

Factor loadings of stable and labile groups compared on Promax Solution. Also given are original Lienert communality estimates, and Promax intercorrelations between factors. A brief description of each test is quoted from Lienert, and also the test's suggested factor composition.

Next, Lienert carried out a multiple factor analysis following Thurstone's (1947) procedure. It was found that eight factors could be extracted from the stable group and only four from the labile group. Communalities were lower for the labile than for the stable group and specific factors were more important for the labile than for the stable group. After rotation, it was found that three factors could be interpreted for the labile and six for the stable group; the latter were said to be closer to Thurstone's primary factors, whereas the former were much more mixed. These figures suggest strongly that children high and low on neuroticism differ very significantly in the way their mental abilities are structured. This conclusion is so important that a thorough critical analysis of the study seems in order.

The first point of criticism is that too little information is given about the analysis to make detailed evaluation possible. The only reference is to Thurstone's book (1947) which contains a number of different methods of analysis, and it is not possible, for instance, to find out just what criteria were used for the extraction of factors or for the interpretability of factors.

Even more disturbing is the failure of the discussion to agree with the results given. Thus, for instance, Lienert says (page 149) that "factor A is a purely verbal factor because it has substantial loadings only in verbal tests." Inspection of Table 5 (b) shows that factor A has the highest loading on a reasoning test, the second highest loading on a word fluency test, the third highest loading on a number test, the fourth highest loading on a space orientation test, and the fifth highest loading on a space orientation test. The sixth highest loading is on a reasoning test. Thus, of the six tests with the highest loading on factor A only one could be interpreted as representing a verbal factor. Factor B is said to be a reasoning factor, having its highest loadings on two tests which, in actual fact, have nearly the lowest loadings on this factor. Altogether, we were unable to make the figures agree with the interpretations, and this must cast doubt on the analysis as a whole, and the conclusions derived by Lienert.

Fortunately, the original matrices were given in the paper, and thus it was possible to carry out a re-calculation based on more modern analytic methods of factor rotation. The method of analysis used by us was Hotelling's principal axes method.

Guttman's (1954) well-known lower bound for the number of common factors indicated the number of factors to be retained. The number is equal to the number of latent roots greater than one in the correlation matrix with unit diagonals. This corresponds identically to Kaiser's (1962) upper-bound for the number of factors with positive generalizability (a term introduced by Cronbach, *et al.* (1963) for the old notion of internal consistency reliability). Three factors were indicated for the stable group and two factors for the labile group. This is in marked contrast to Lienert's solution in which eight factors are retained for the stable group and four for the labile group. The reason for this discrepancy is difficult to assess since Lienert does not indicate his criteria for this decision. Probably, it is largely due to the inefficiency of the centroid method relative to the principal axes method. But, since we have no indication as to the reflection procedures used in the centroid analyses which Lienert presents, the relative efficiency cannot adequately be assessed.

With the number of factors thus fixed, the communalities were estimated by the now standard procedure of iteration by refactoring (Harman, 1960). The method of principal axes was used and after fifteen cycles, all communalities and converged to three decimal places (though most had converged to four or five places). The final communality estimates are presented in Table 1. For comparison, we also present Lienert's estimates in the same table.

The marked tendency towards very much lower communalities for the labile group which Lienert notes is not so apparent in the present analysis. Since Lienert does not indicate his basis for estimation, the reason for the discrepancy cannot be evaluated.

With the communality estimates thus determined, and the number of factors fixed as before the factor loadings were computed for each matrix by the method of principal axes. Kaiser's (1956, 1958) Varimax procedure for analytical rotation to orthogonal simple structure was applied to the principal axes matrices. The Promax (Hendrickson and White, 1964) procedure for analytic rotation to oblique simple structure was applied to the Varimax solutions. The oblique factor loadings for each matrix appear in Table 1, along with the intercorrelations among the primary factors and the test communalities. The principal axes loadings and the intermediate Varimax loadings are not presented here but all relevant matrices are available at the Institute of Psychiatry.

Upon inspecting the patterns of loadings presented in this table, one is not particularly impressed by the clear and unambiguous interpretability of the resultant factors. Indeed, the crisp, clear simple structure usually associated with P.M.A. material is nowhere to be seen. However, oblique rotation has cleaned up the simple structure considerably and tentative hypotheses may be put forth for at least some of the factors. For convenience of reference,  $S_1$ ,  $S_2$ , and  $S_3$  will indicate the respective factors for the stable group, while  $L_1$ ,  $L_2$  will indicate those for the labile group.



Factors  $S_2$  and  $L_1$ , seem primarily to involve the use of words. In each case, the four tests with highest loadings were postulated as measures of either the Verbal Comprehension factor (V) or the Word Fluency factor (W). Factor  $S_3$  has its three highest loadings on tests postulated as measures of the Spatial Relations (S) factor and no other loadings exceed 0.40. Factors  $S_1$ , and  $L_2$  appear to be rather complex. No very simple interpretation is suggested although the pattern of loadings appear to be rather similar for the two factors. Each is loaded by tests hypothesized as measures of Reasoning (R), Spatial Relations (S) and Closure (C). Additionally,  $S_1$  has moderate loadings on putative measures of Verbal Comprehension (V) and Word Fluency.

Our own solution, while differing considerably from Lienert's, does suggest that his main contention is indeed borne out by his data; the stable group has a more clearly marked structure in the cognitive test field than has the labile group. Three significant factors in the stable group are opposed to two significant factors in the labile group, and as the same standards of selection were employed at all stages, there seems little reason to doubt that these differences are real ones rather than being statistical artefacts. It will, of course, be necessary for this work to be repeated, preferably with a larger selection of tests, before the revolutionary implications of Lienert's work can be accepted; nevertheless, it would seem likely that personality and intelligence test performance are indeed more closely imbricated than has hitherto been thought likely.

It will have been noted that there is a curious symmetry in the results obtained by Lienert, and those obtained by Shure and Rogers. High ability subjects show higher variance of the N factor than do low ability subjects. High stability subjects show greater organization of abilities than do labile subjects. It would almost appear as if greater stability and ability, respectively, went with greater degrees of organization of ability and stability. It is much too early to speculate about the possible meaning and causes of these relations; much further research is required before the facts themselves are adequately established to call for explanatory hypotheses. Nevertheless, the theory of linear independence between cognitive and non-cognitive factors may soon have to be supplemented by one stressing non-linear dependence and interrelation.

## REFERENCES.

- CATTELL, R. B. (1963). The theory of fluid and crystallized intelligence: a crucial experiment. *J. Educ. Psychol.*, 54, 1-22.
- CRONBACH, L. J., RAJARATNAM, NAGESWARI, and GLESER, GOLDINE C. (1963). Theory of generalizability: a liberalization of reliability theory. *Brit. J. Stat. Psychol.*, 16, 137-163.
- EYSENCK, H. J. (1953). *The Scientific Study of Personality*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1957). *The Dynamics of Anxiety and Hysteria*. London: Routledge and Kegan Paul.
- EYSENCK, H. J. (1959). Personality and problem solving. *Psychol. Rep.*, 5, 592.
- EYSENCK, H. J. (1960). *The Structure of Human Personality*. London: Methuen.
- EYSENCK, H. J. (1960). *The Structure of Human Personality*. London: Pergamon.
- EYSENCK, H. J. (Ed.) (1964). *Experiments in Behaviour Therapy*. London: Pergamon Press.
- FURNEAUX, W. D. (1962). The psychologist and the university. *Univ. Quart.*, 33-47.
- GUTTMAN, L. (1954). Some necessary conditions for common-factor analysis. *Psychometrika*, 19, 149-161.
- HARMAN, H. (1960). *Modern Factor Analysis*. Chicago: University of Chicago Press.
- HENDRICKSON, A. E., and WHITE, P. O. (1964). Promax: a quick method for rotation to oblique simple structure. *Brit. J. Stat. Psychol.*, 17.
- HORN, W. (1962a). *L-P-S Leistungsprüfsystem*. Göttingen: Hogrefe.
- HORN, W. (1962b). *S-O-F Stabilitäts- und Offenheitsfragebogen*. Unpublished Ph.D. dissertation. University of California.
- KAISER, H. F. (1956). *The Varimax method of Factor Analysis*. Unpublished Ph.D. dissertation. University of California.
- KAISER, H. F. (1958). The Varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- KAISER, H. F. (1962). Image analysis. In: *Problems in Measuring Change*. Ed. C. W. Harris. University of Wisconsin Press, Madison.



- LIENERT, G. A. (1963). Die Faktorenstruktur der Intelligenz als Funktion des Neurotizismus. *Ztschr. f. Exp. Angew. Psychol.*, 10, 140-159.
- LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational attainment. *Brit. J. Educ. Psychol.*, 31, 194-203.
- SHURE, G. H., and ROGERS, M. S. (1963). Personality factor stability for three ability levels. *J. Psychol.*, 55, 445-456.
- THURSTONE, L. L. (1938). Primary mental abilities. *Psychometric Monogr. No. 1*. Chicago: University of Chicago Press.
- THURSTONE, L. L. (1947). *Multiple Factor Analysis*. Chicago: University of Chicago Press.
- VERNON, P. E. (1958). *The Structure of Human Abilities*. London: Methuen.

## PUBERTY AND TEST PERFORMANCE: A FURTHER REPORT

By J. D. NISBET, R. ILLSLEY, A. E. SUTHERLAND AND M. J. DOUSE

(University of Aberdeen)

**SUMMARY.** We reported previously from successive testing of an age-group of girls at ages 7, 9, 11 and 13, that those who reached puberty before age 12:3 had a slight but consistent superiority in average intelligence test score. A further testing of a selected sample from this group at age 16 suggests that this superiority diminishes as the late developers reach puberty.

### I.—INTRODUCTION.

In a previous paper (Nisbet and Illsley, 1963) we reported on the relation of early onset of puberty to test performance at age 11, on the basis of a follow-up of a year-group of girls from age 7 to age 13. The results presented there agree closely with the findings of Douglas's long-term study under the auspices of the Population Investigation Committee, reported in *The Home and the School* (1964: Chapter X). Both studies showed that girls who reached menarche relatively early obtained on average slightly higher scores at age 11; but this slight superiority was not associated with onset of puberty, being as marked at age 7 or 8 as at the age of 11.

It is clearly important to extend the age-range of our investigation to cover the later stages of adolescence also. To quote Douglas (*op. cit.*, page 82): "Although the girls who mature early have superior measured ability at 11 years, there is no certainty that they retain this advantage in later school life . . . It is important to know what is happening. If the early developers retain their superiority in performance throughout their school life, there is no reason to allow for maturity when selecting for secondary schools, but if they lose their initial advantage and the late developers catch up, we are selecting inefficiently and should do something about it."

The differences in mean score between early and late maturers are relatively small, compared with the spread of scores in both groups which extends over the full range from highest to lowest. We should not, therefore, exaggerate the importance of the differences for selection at age 11. However, the finding is of interest for the understanding of influences on mental development as measured by tests of reasoning.

The extension of inquiry to include the later stages of puberty has been done in Aberdeen, though only on a selected sample of the original group, namely, those who stayed on at school after age 15 in the senior secondary schools. In our previous study, which stopped at age 13, we were able to identify the early and very early maturers. In this extension, we were able to identify also the late maturers separately from the median group.

### II.—SUBJECTS AND TESTS.

The population of the original inquiry was defined as those girls who were in the second year of secondary schools in Aberdeen in December, 1960. In June, 1963, when the girls were age 16, those in the senior secondary schools were given Moray House Adult Test I. They had done Moray House Advanced Test 10 two-and-a-half years previously, and for the majority age at menarche was already



known. Letters were sent to the mothers of the eighty girls from this group who had not reached menarche at the time of the early inquiry, and almost all replied, giving the information requested.

The total number of girls for whom scores at ages 13 and 16 and age at menarche were available, was 273. These were classified as: very early maturers (menarche before age 12:3, 53 girls); early maturers (menarche between 12:3 and 12:11, 62 girls); median group (menarche between 13:0 and 13:8, 86 girls); and late maturers (menarche after 13:8, 72 girls). Menarche occurred before age 11:0 in 1 per cent.; before age 12:0 in 13 per cent.; before age 13:0 in 42 per cent.; after age 14:0 in 22 per cent.; and after age 15:0 in 10 per cent.

### III.—RESULTS.

In our previous report, all scores were restandardised on the population studied to remove inequalities in test standardisation. With Moray House Adult Test I at age 16, such restandardisation was not possible since only a selected group was tested. It was, therefore, necessary to use the published norms for the Adult Test, and to express the analysis in terms of deviations from the group average, corrected for differences in standard deviation. For Moray House Advanced Test 10, at age 13, scores on the revised standardisation were used as these were recorded on the punched cards.

Table 1 gives the results of the comparison of mean test scores of the very early, early, median and late maturers at ages 13 and 16. The mean score of the whole group at age 13 on the revised norms was 117.82 (standard deviation 9.68); and at age 16 on the published norms, 125.26 (standard deviation 10.87).

TABLE 1  
MEAN SCORES AT AGES 13 AND 16 BY AGE AT MENARCHE.

Menarche category	Very early	Early	Median	Late
At Age 13:				
Mean score .....	121.62	117.81	116.94	116.10
Deviation from mean as z-score ..	+0.393	-0.001	-0.091	-0.178
At age 16:				
Mean score .....	126.81	124.66	125.26	124.64
Deviation from mean as z-score ..	+0.143	-0.055	0.	-0.057
Number of cases .....	53	62	86	72

The results show that the earlier average superiority of the very early maturers has been reduced substantially at the later age, and that late maturers have made up much of their previous inferiority in average test scores. At age 13, the mean score of the very early maturers was significantly different from that of the remainder; there was a significant change in mean score between ages 13 and 16; and at age 16, their mean score was not significantly different from that of the others. The means in the late group showed no significant differences throughout, though the results are consistent with the hypothesis that they are catching up on the others. The correlation between scores at 13 and 16 in this selected group was +0.701.

ACKNOWLEDGEMENTS.—We acknowledge with thanks the assistance of Dr. D. Younie, Mr. J. R. Clark, Miss E. Currie, Miss D. Kidd, Miss M. McNab, Mr. J. Goldie, and Mr. B. Macdonald. The cost involved was met from a grant from the Scottish Council for Research in Education.

### IV.—REFERENCES.

- DOUGLAS, J. W. B. (1964). *The Home and the School*. London: Macgibbon and Kee.  
 NISBET, J. D., and ILLSLEY R. (1963). The influence of early puberty on test performance at age 11. *Brit. J. Educ. Psychol.*, 33, 169-176.

## CRITICAL COMMENT

### COMMENTS ON THE ARTICLE BY MAGNE AND PARKNAS

By M. D. VERNON

(*The University, Reading*)

I was much interested in the paper published in Vol. 33, Part 3, of the *British Journal of Educational Psychology*, by Olof Magne and Lenart Parknas, on "The learning effect of pictures." I think, however, that there are certain points which must be borne in mind in interpreting the results of experiments on the use of pictures in teaching. The first is the nature of the material illustrated. From pictures, the over-all shapes and more obvious characteristics of objects can be effectively perceived, and remembered for discrimination later, as is shown by the flower pictures in the first two experiments described in this paper. It is more doubtful whether the detailed inter-relations of parts within the whole can be retained for subsequent identification and further conceptualization without verbal labelling. Experimental evidence would suggest not.

When pictures are used as illustrations to a text or a lesson, the case is somewhat different. Some parts of the content, for instance, the appearance of scenes in foreign countries, may be presented most effectively by pictorial means, and subsequent recall, especially of factual items shown in the pictures, will be correspondingly facilitated. Pictures which are merely illustrations of what is clearly described in the text are less useful, except in so far as they arouse interest, which may often occur with young children. But, if people are required to remember the ideas or the general gist of a text or lesson rather than specific details, pictures may be distracting since they give too much emphasis to factual detail and too little to the general argument.

Thus, it is unwise to suppose that there can be any general answer to the question, "Is pictorial material useful in teaching?" What seems to be needed is a systematic exploration of the situation and material presented, the subsequent application of what is learnt, and the age and interests of the learner.



## CRITICAL NOTICE

*Year Book of Education, 1963: The Education and Training of Teachers.*  
Edited by G. Z. F. Bereday and J. A. Lauwerys. Evans Bros., pp. xiv + 578, 63s.

By STEPHEN WISEMAN

The *Year Book* is nothing if not topical. Last year it gave us the *Gifted Child*; for the year of Robbins we have the *Training of Teachers*. As usual, it is a weighty production: nearly 600 pages, divided into fifty-five chapters from fifty-three contributors. This reviewer hopes that the rise in the number of contributors (from last year's forty-one) does not signify a trend: the pursuit of quality and relevance by the editors ought undoubtedly to produce a movement in the opposite direction.

The book is arranged in four sections: I—*Historical and Theoretical Studies*; II—*The Organisation of Teacher Training: Area Studies*; III—*Social and Economic Problems*; IV—*New Experiments in Teacher Education*. The second section, with twenty-two chapters covering eighteen different countries, is more than twice as long as each of the other sections, and in weight at least forms the core of the volume.

In a short introduction the editors confess that the over-all view of teacher training "is not heartening." It is still handicapped by "the notion that the teaching profession is the first step on the white-collar ladder for the ambitious members of the working classes"; training institutions lack status and prestige, having "their faculties overwhelmingly selected from working practitioners in the schools themselves without due regard to the type of qualities appropriate to the post of university professor"; teachers, old and young, exhibit a "paralysing conservatism." This inspissated gloom is only lightened by the editors' discernment of an "almost increasing awareness of the need for a 'thorough re-appraisal,' and an 'almost universal tendency to upgrade all teacher training institutions to university status.'"

The volume proper begins with a sensible and balanced general introduction by Holmes who considers the changing concept of the ideal teacher. The education of teachers is bedevilled by that familiar paradox, that education must conserve the culture of the past, and yet must prepare for, and initiate, progress in the future. Many of our difficulties and tensions are built-in as a result of this inevitable ambivalence. Holmes touches on the status of the profession, and the active Lieberman's side in demanding a "research minded" profession, and the active participation of teachers in developing (and applying) a science of education. Only in this way can the teaching profession "rest on a new basis of authority." This needs to be said much more often, and it is noticeable that many contributors to this volume implicitly reject such a point of view.

The most interesting chapter in Section I for readers of this *Journal* is *Psychological Theories and Teacher Training* by Professor Tibble. It is historical in form, leading us from James and Sully up to the present day. Or nearly up to the present day, since there is a curious air of *fin de demi-siècle* about it—only one sentence on Piaget, and no mention of programmed learning, for example. Tibble concerns himself with the impact of psychological theory on the whole of the college course, and not merely on the content of the psychology course. He rightly emphasizes the way in which the philosophy of Dewey and the psychology of Freud, responsible for the change from teacher-centred to child-centred methods with young children, spilled over into the colleges to produce 'student-centred' programmes and methods. Students "by learning to understand children could come to know themselves." Tibble does not recognize the danger of the lengths to which this has gone, for now the emphasis is reversed: students by learning to understand themselves, can come to understand children—and if the second objective is not attained, then we can still be satisfied with the first. He quotes with approbation J. L. Henderson's pontification "the core of a curriculum for teacher training should be the skilful elucidation of the nature of human relationships through the experience



ing of them. What the students learn about them, and even more how they learn about them, will largely determine their subsequent treatment of children." Less than half the teachers of psychology in the colleges have any formal qualifications in psychology; the rest of the staff (the majority) are all untutored in the subject. The contemplation of such lecturers applying themselves enthusiastically to student therapy and the removal of emotional blockages is far from reassuring. The experiences of teacher-education in the U.S.A. ought to make the colleges pause. The status and prestige of 'education' as a subject there has probably never been lower. Criticisms of the 'anti-intellectualism' of professors of education and of education textbooks and courses are now so frequent that it is becoming more and more difficult to shrug them off as the splenetic comments of educational backwoodsmen. Such writers as Berzun and Koerner cannot be entirely ignored. Tibble himself proposes the remedy: the problem for the educator of teachers is "to select from the increasing body of solid psychological research . . . those studies which bear most relevantly on the problems he meets." Holmes and Lieberman are better guides than Henderson.

King's chapter on the curriculum of British teacher training courses exhibits some of the same tender-mindedness as Tibble, particularly when he comes to discuss the perennial battle between technique and content. "We recognise," he says, "that civilization and education depend upon attitudes. These derive from our practically sharing in complementary responsibility which educates 'as we go'—teaching content and technique inseparably, and assessing both of them by criteria of responsible personal involvement in the dynamic social ecology of our progress." One wishes that Dr. King—in equally impeccable "educant," to use Koerner's graphic term—would apply this doctrine to the student training for infant schools, and taking mathematics or English Literature as a Main subject. It is a far cry from differential equations to Cuisenaire rods, or from the Initial Teaching Alphabet to T. S. Eliot.

Section II, *Area Studies*, opens with an excellent introduction by Holmes, perceptive and balanced. The following twenty-two chapters are, inevitably, uneven in quality and relevance. The chapter by Baron on England and Wales, is competent and factual, but—inevitably—already out of date in the post-Robbins period. Kellmer Pringle, on the training for remedial work, leans heavily, and understandably, on the Birmingham pattern, while Rose Bruford, on the training of teachers for speech and drama is, unfortunately, completely non-critical. Elsbree's chapter on the U.S.A. brings us (after nearly 200 pages) to the first mention of research, and a recognition of "the relatively low standards which obtain in many of the 1,200 institutions which prepare teachers." The reasons for this are low salaries and low prestige, producing acute staffing problems. The perception of a two-way interaction here is nowhere obvious. One of the best chapters in this section is by Katz, on Canada; and one of the most interesting by Dottrens, on Switzerland. Katz gives us much more than an account of the Canadian scene—itsself exhibiting many contrasts—but provides a critical survey of some of the basic questions in all countries. The main interest in the Swiss contribution is an account of college selection methods in Geneva, which include giving a talk to a class of children, and teaching an arithmetic problem. "Candidates range from those who cannot put a foot wrong and who capture the interest of both pupils and examiners in a masterly fashion, to the unfortunate wretch who comes into the classroom as though he were entering a lion's cage, rushes to the teacher's desk, pours out what he has prepared without once daring to raise his eyes and look at the children in front of him, and then rushes out without a backward glance at the point of collapse." Unfortunately, there is no suggestion that any validation studies have been performed, nor any comparison with more traditional selection methods. Other chapters in this section provide useful information about teacher training in other countries and, occasionally, support for our national stereotypes. Who could doubt, for example, that the two following quotations come from France? "What counts most in teaching in the vocation of the educator is faith and conscientiousness"; and, "life offers adequate solutions to the problems it poses, if they are examined with an open mind."



On the whole, Section II is disappointing, mainly perhaps because of its overwhelming weight on Europe and the U.S.A. Only two or three of its chapters deal with less fortunate countries. The excellent introduction by Wayland to Section III underlines the paucity of II and the limited view of the majority of its contributors, with an almost complete absence of data on such matters as language problems, competition with other professions, the urban-rural dichotomy, ideological struggles and sub-cultural groups. One wishes that the writers in Section III itself could have had a pre-view of Wayland's short chapter: the scope of their contributions might have been wider. Kob, from Western Germany, gives us an analysis of the social status of the teaching profession, which leads him to postulate deep inferiority feelings on the part of the teachers. His thesis, however, is less than persuasive and his analysis is too superficial. It suffers, no doubt, by his ill-luck in coming just before Langeveld's penetrating and stimulating analysis of the psychology of teachers and the teaching profession. He begins by defining the excellent teacher as "one who can first of all teach in such a way that the young understand his teaching well and follow his lessons with a satisfaction which justifies their absence from play for themselves or from work for their parents . . . To be a teacher is a function as well as a profession." He emphasizes the individual differences among teachers (when is a research going to concern itself with these?) and insists that "every teacher has a history, he has to make use of it in a positive way." All of Langeveld's twenty pages are stimulating, and ask for quotation: space limits one to four only. "What a school system needs is not only an entrance for the teacher but also an emergency exit." "If teachers are trained in such a way that they become specialists who can only survive in a given school system, they will be resolute opponents of any kind of change and will cling to that system to their last breath." "Every real educational situation offers an optimum possibility and a maximum risk, and whosoever does not like this must choose a safer occupation, for instance, that of schoolmaster!" "If man could live to become 5,000 years old, it would always be possible to hear from a teacher that he has been at it 4,975 years, so what he does must be right." An analysis of characteristics leading (a) to unfavourable, and (b) to favourable occupational adjustment, is the final outcome of this wise and witty chapter, and is so interesting and suggestive that one feels more than usually frustrated at the absence of any references to experimental work which might support Langeveld's hypotheses.

A chapter by Sandven from Norway on predicting teachers' attitudes raises one's hopes, but it turns out to be another example of that *bête noir* of this reviewer at least, an early report of a research in progress, telling us what they hope to demonstrate. Marklund reports on Swedish teachers' attitudes towards a revolutionary comprehensive school system and provides not only experimental results, but tests of significance—in which respect his chapter stands out like a good deed in a naughty world—but the final outcome, though interesting, is depressing. It provides yet another example of the "paralyzing conservatism" noted by the editors, and Marklund's last sad sentence reads: "To a far greater extent than was foreseen at the outset, the school reform has proved to be a teacher problem." A chapter on teacher recruitment in France shows the familiar U.K. pattern of bulge and trend, with a severe forecast shortage in both lycées and universities, in spite of emergency measures.

The reader, by this time, has arrived at p. 476, and he turns to Section IV, *New Experiments in Teacher Education*, with relief and with anticipation. I regret to report that he is unlikely to experience the first or to satisfy the second. The first chapter, by Jordan, on new experiments in the U.K., endeavours to satisfy the chapter title by describing the programme at a single training college, his own. The fact that "the work is not overshadowed by written examinations: the dead hand of the three-hour written paper with its inadequate challenge, its insufficient probing, its chancy nature, has largely been removed," may account for this glaring example of not answering the question set. A rather pedestrian description of the foundation of Brinsford Lodge, the Malayan Teachers' Training College in the Midlands, is followed by four chapters on developments in the U.S.A. That by Eurich, on the role of the American Foundations in stimulating research and

development, is a useful, sober and sensible statement, while Cartwright's survey of some of the outcomes of post-Sputnik anxieties illustrates the "awareness of the need for a thorough re-appraisal" noted by the editors. The other two chapters, on the New Horizons Project, and the Certification programme, are highly 'professional,' in the narrowest sense of the term. There is little mention of the criticisms mounted against N.C.A.T.E. by liberal arts professors (as might be expected when accreditation concerns itself only with the professional side of the teachers' courses), nor any mention of the famous rejection of Carleton College, in 1960, followed by a reversal of the decision in 1962. The chapter on state certification gives us little indication of the range of differences among states (a table here would have been invaluable) or of recent trends in, say, the proportion of time demanded for professional as against academic studies. These chapters are much less informative for the trans-atlantic reader than they might have been. An interesting chapter by Minogue, on New Zealand, discusses the idea of a B.Ed. degree, and in the process gives us a useful—and, for U.K. readers, topical—discussion of the cultural *versus* the professional study of education.

In conducting a review of a book such as this, one is inescapably reminded of the Gestalt insistence that the whole is more than the sum of its parts. An attempt to indicate the coverage provided by so many and so diverse chapters, leads me into an over-emphasis, perhaps, on criticism of detail. If this leads the reader to an unfavourable view of the book as a whole., this would be manifestly wrong and unjust. I will not use the common cliché and suggest that this volume is a 'must' for all engaged in teacher training, but I believe that none could read it without profit. Its value lies not only in what it says, but in what it forces us to think about, and practices and programmes it forces us to re-examine. Its major omission—and it is a serious one—is that of research in teacher training. The results from this have so far—been meagre and unrewarding. But the readers of this volume deserve to have some review of research, some suggestions for future action, and some comment on promising lines of attack—for example, the work of Pace and Stern on the College Characteristic Index. And I cannot forebear to make one final comment, in the form of a plea to the editors to re-consider their stated policy to "exercise no control over what the authors wish to say. We select our contributors with an eye to representation of different countries and topics as well as to their personal calibre. Once selected, we exercise only minimal editorial prerogatives over standards of presentation." One of the most effective instruments of research and development in education is—a blue pencil.

STEPHEN WISEMAN



## BOOK REVIEWS

DAVID, H. A. (1963). *The Method of Paired Comparisons*. London: Charles Griffin, pp. 124, 28s.

This is another excellent volume in the series of Griffin's statistical monographs and courses. It deals with a method of analysis which, although restricted in use, may still be of considerable service to experimental psychologists.

The basic experimental unit dealt with is the comparison of two objects (or items, treatment, etc.) by a judge who has to choose one of them, and, when more than two such objects have to be considered, a choice has to be made from every possible pair. Obviously, it is only when objective measures are precluded, and when simple ranking or grading devices are impracticable, that such a situation is usually resorted to. Thus, one possibility to be encountered at the outset is that of circular triads (i.e., A is preferred to B, and B to C, yet, nevertheless, C is preferred to A). But, as the author rightly points out, it is not a defect but a valuable feature of the method of paired comparisons that these seeming contradictions are allowed to manifest themselves. The complete absence of such contradictions (due possibly to a judge's good memory) implies that the comparisons have degenerated into rankings. A possible answer lies in designing the experiment so that a single judge makes only part of the possible comparisons. The discussion later in the book on incomplete and linked-paired comparison designs is, from this point of view, particularly appropriate.

Before discussing design, however, the basic probability models—such as the *stochastic transitive*, the *strong stochastic transitive* and *linear* models (including the Thurstone-Mosteller, Bradley-Terry and Bock's three component model)—are reviewed, and a variety of significance tests (many of which parallel the better-known tests for differences of treatment means in an analysis of variance) are described. There are many interesting applications—including for instance how in a cup-tie competition one can best replicate to be more confident that the *second-best* side has been correctly selected—and a final chapter deals with a miscellany which includes the selection of judges, of consumer panels, and the treatment of ties. Most of the chapters conclude with exercises for the serious student, and a compact list of references and statistical tables is appended. In short, this will undoubtedly prove a most valuable book to every psychologist compelled, from the nature of his research interests, to work in the non-parametric field.

D. G. LEWIS.

DRILLIEN, C. M. (1963). *The Growth and Development of the Prematurely Born Infant*. Edinburgh: Livingstone, pp. vi+376, 50s.

This is an account by a woman paediatrician of a follow-up study from birth of a large sample of premature children and controls. For their first five years, periodic visits were made to the homes during which were recorded, besides the health and development of the children, the maternal attitudes, standard of care and degree of stress to which the child was subjected by deprivation or insecurity. Six developmental assessments were made between 6 months and 4 years. At 5 years, the Edinburgh School Psychological Service administered the Terman-Merrill, and assessments of social adjustment were made around the 7th year. Methodologically, the study seems unexceptionable. The chief results are given against a grading of cultural grouping, thus overcoming the pitfall that more premature children are born in the poorer sections of the working class.



The findings are of prime importance for the understanding of the determinants of infant behaviour. They demonstrate, first, that congenital factors influence both I.Q. and social adjustment and, second, that the nature of the handicap which they induce is a greater vulnerability to post-natal stress. A revealing series of histograms on p. 199 show the usual social-class differentials for I.Q., but there is a remarkably constant difference of 20-21 points between the mean scores of the small prematures and maturely born within each social class. Unfortunately, although exceptionally full data are usually given, in this case the author has omitted to give the figures in tabular form. In view of their interest to psychologists, it is hoped that this omission will be made good in a later edition.

The phenomenon of vulnerability is very clearly brought out for social adjustment (p. 254). Where the only adverse factor is prematurity, there is only a moderate difference compared with the maturely born. But, in stressful family-situations, the prematures more than double their 'maladjustment' scores while those of the matures rise by only 50 per cent. Pregnancy and birth complications add cumulatively to the degree of behaviour disturbance at 7 years. Even more remarkable was that, to quote the author, "three-quarters of those mature controls who were subjected to severe familial stress in their early years appeared to be emotionally stable, and less than 1 in 10 showed borderline to obvious maladjustment" (p. 251). Where congenital factors were present, the proportion maladjusted rose to one-half. Although anxious mothering produced more pre-school behaviour problems, these were not reflected in adjustment to school.

As this sample is followed through the school years, the results should be of increasing interest. There will henceforth be no excuse for a failure to take account of congenital factors in child-deprivation studies.

D. H. STOTT.

HERRICK, VIRGIL E. (Ed.) (1963). *New Horizons for Research in Handwriting*. University of Wisconsin Press: Madison. pp. 276, 5\$.

This book is a report of an invitational conference on research in handwriting. Lest there be any misunderstanding, it should be said at the outset that handwriting is looked at from the point of view of a skill to be acquired rather than as a means of expressing character.

The core of the book consists of eight papers which range from reports of large-scale surveys via attempts at setting up handwriting scales to laboratory studies of handwriting as a sensory-motor skill. Each of these papers is followed by comments from the discussants.

The book is of general interest because it is an account of an attempt at a 'global' attack on a real life problem. The reader who has encountered similar situations will now meet many familiar arguments; and he who has not could do worse than to take this for a model: should one examine the existing system in all its aspects, or concentrate on a projected system? Why handwriting at all? Who uses it and for what purpose? If one does tackle the existing system, where should one start, with the finished product achieved by adults or with the teaching of children? Should one evaluate handwriting on grounds of artistic merit or on ease of comprehension? How far should the possible practical needs of education constrain research at the expense of fundamental work? One of the pleasures of the book is to find these and other points raised, more often by discussants perhaps than by the contributors, discussed and happily left unresolved.

Handwriting, like speech and driving, is an everyday skill, the subject much more frequently of speculation than of research. The information gleaned by the survey of the art as practised in schools and recommended by commercial systems of handwriting is presented by Herrick and Okada. It represents the basis for a whole programme of research. The information is impressive though neither the remaining papers nor much of the final recommendations seem to draw upon it. As one discussant emphasized: "It indicates that we cannot use a report about what is 'most practised' to tell us about what is 'best practised'."



The paper, by Harris and Rarick, once again shows up the difficulties of measuring skill by inference, in this case by recording pressures during writing. In the following paper, K. U. Smith, expounds neorogeometry as applied to handwriting. It is an odd experience to find somebody clearly appreciating contemporary concepts of communication systems engineering, but insisting on using something like late nineteenth century language when describing them. It is then pure delight to find handwriting performance studied by distorting and disrupting visual feedback by virtue of closed circuit television and video tape, even though the data presented here relate to few subjects only. Harris and Herrick, in a paper on children's perception of the handwriting task, show up the apparent inability of children to assess their own performance relative to a standard set by the teacher. Herrick and Erlebacher go into considerable detail to describe their methodology for building a legibility scale. Fascinating and forbidding when one considers the labours which must have gone to obtain a result of very limited usefulness. And yet, if it is indeed necessary to have such scales, this paper is a solid piece of workmanship.

Whether or not there is, indeed, a need to know more about handwriting in the United States at present, this book may help to draw attention to the use of handwriting as a more or less accomplished human skill to be studied in its own right. One may or may not agree with much of K. U. Smith's theorising, but surely he is right in his claim that so far orthodox experimental psychology tends to 'observe and blanket' problems of complex behaviour organization. It has done so because, with a few exceptions, the tasks chosen for laboratory study have been inhumanly simple. Handwriting, as the book shows, is not. J. A. LEONARD

HILL, WINFRED F. (1963). *Learning: A Survey of Psychological Interpretations*. San Francisco: Chandler Publishing Co., and London: Methuen pp. 227.

From the Preface: "Systematic interpretations of learning have an important place both in the science of psychology and in the application of psychology to education. This introductory survey of contemporary learning theories is designed to provide a fairly elementary but, nonetheless, solid account of this topic for students in the psychology of learning and in educational psychology."

The application of psychology to education is not exploited in the text, but the intention stated in the second of the quoted sentences is more than adequately fulfilled. The material is solid, but far from stodgy. The text is eminently readable, enhanced by a sufficiency of examples, and, for good measure, enlivened by a quietly humorous style and the occasional apt anecdote. Most books on psychology are too long; your reviewer was sorry this book was not longer.

There are neatly condensed accounts of the work of Watson, Guthrie, Thorndike, Skinner, Miller, Lewis, Tolman, Hull, Spence, Mowrer, Estes, and of some Gestalt views. A chapter deals with some 'miniature' systems, concerned with more restricted fields of inquiry. The book ends with some thoughts on the shape of theory to come.

The summaries of the theories seem to be as fair to their authors as summaries could be. Also, in presenting so many theories in so short a space, the author has contrived to avoid the 'and-and-and' style deplored by good Gestaltists, but found in too many texts, and instead to create an organisation with good interrelations. This has been done by his classifying the theories, first of all, in terms of 'connectionist' and 'cognitive'; by his showing how theories of the one form differ among themselves; and by his relating theories of the one form to theories of the other. The theories are further interrelated by the author's view that the two forms of theory are gradually moving together.

Your reviewer, as a teacher of psychology, would be very much in favour of students' being set to read this book before being expected to make a detailed study of particular learning theories or of the classic anthologies. It will, I think, help the student to put the particular theory in context, and to whet his intellectual appetite. JAMES M. THYNE.



RICHMOND, K. (1963). *Culture and General Education*. London: Methuen, pp. 184, 21s.

In this book the author publishes several 'tests of culture,' of which two produced considerable interest when they appeared in the *Times Educational Supplement*. With them there is a lively and, at times, even racy, comment which—together with its subject—will guarantee the book a good sale. There is also a heterogeneous collection of tests results, some of which the author admits to be from groups which are by no means representative of their parent population. Other norms are more representative, and there is an interesting comparison of the scores of university students from the Faculties of Arts, Science, Applied Science and Law. A similar comparison is made of the scores of university dons, undergraduates and sixth formers. A noteworthy feature of the tests is the use which is made of an apposition of the sciences and the arts; an individual's score, therefore, gives some indication of the extent to which his interests in these fields are well balanced.

The author admits that there are many weaknesses in the book, but claims that this is inevitable in a virtually new field. While this is partly true, and one sees that some of the difficulties could not easily have been met, yet the publication of such a work throws upon the author the responsibility of ensuring that his readers will not get a misleading impression. Though most psychologist readers will keep in mind the author's cautionary statements and apply them throughout the book, it is doubtful whether the general public will do so. There are also occasions when the wording gives an impression which Mr. Richmond may not have intended. He writes phrases such as "the results show that invariably (the test) discriminates between good schools and those which are not so good" (p. 137), and (when comparing independent and direct grant schools on the one hand, with state schools on the other), "there is no escaping the conclusion that the process of levelling up in state-provided schools has a long way to go before our pious talk about equality of opportunity can escape the charge of being so much claptrap." These phrases seem to indicate that the author believes that the differences between the state schools and the others in culture score is caused by better teaching in the latter. Though it is generally supposed that there is this difference in the quality of teaching, this is a questionable assumption. In the present context much more likely causes are differences of social class between the schools and the very high rate of special selection in the direct grant schools. It is also possible that this social class factor may be responsible—in part at least—for some of the differences in average score between various faculties.

There is one other point. Though the author shows that he is aware that the results he obtains from his tests are in part a function of what he puts into them, he occasionally forgets this in his interpretations. In particular, he finds that history specialists do rather badly, and it seems to the reviewer that this might be because the test includes more material which is directly from the subjects studied by science students and by students of literature, but not so much which is directly from the type of history studied in universities.

In spite of the above remarks, the author is to be congratulated on giving to education in this country a useful continuation, though from another angle, of the pioneer work done by Oliver of Manchester. As the author himself says, more good may come out of the tests than is apparent at the moment; meanwhile, the reader is interested and stimulated.

R. R. DALE.

SUNDBERG, N. D., and TYLER, L. E. (1963). *Clinical Psychology: An Introduction to Research and Practice*. London: Methuen, pp. xviii+564, 50s.

There are probably few areas of psychology where dogmatic positions are so frequently adopted as in the clinical field, where workers adhering to widely differing theories can scarcely communicate with one another without losing their tempers, or at least writing each other off as unworthy.



In this book, the authors take a long cool look at the problems, strategies, methodologies and attitudes of clinical psychologists, offering a number of sage and well-balanced suggestions. The subject-matter is not so much the content of clinical psychology in terms of tests and experimental techniques, but rather an analysis of the sundry behaviours of clinicians in a variety of situations. The reader will find detailed discussion of neither psychoanalytical concepts nor neo-behaviourist systems. He should, however, gain from the experience of clinicians who have obviously read widely in both these and other fields. The book is sub-divided into four sections. The first presents an introduction and conceptual framework; the second, on psychological assessment, covers such topics as the nature of clinical assessment, interviewing, the use of tests, assessing development in life situations, interpretation, communication of findings and research. A major section is devoted to psychological approaches to treatment, in which a clear account of various methods of psychotherapy is presented, together with a widely-cast bibliography. In a final section, the development of clinical psychology itself is reviewed, and a number of thorny problems faced, including that of the relationship between psychologists and psychiatrists. A sensible feature is the presentation of brief summaries of important articles and books appearing at the conclusion of each chapter, in addition to a comprehensive bibliography.

In criticism, this reviewer finds the position adopted somewhat too eclectic; the section on theory is both too brief and insufficiently critical; the authors fail to present an evaluation of research findings in specific areas. Nevertheless, the book may be recommended to students, clinical psychologists and social workers as a useful addition to textbooks of abnormal psychology.

A. D. B. CLARKE.

THYNE, J. M. (1963). *The Psychology of Learning and Techniques of Teaching*. London: University of London Press, pp. 240, 18s.

Too frequently elementary texts in educational psychology are re-hashes of previous books, brought up to date and slanted in accordance with the idiosyncracies of their authors. This is not so with this book, which is a fresh approach to the study of the rationale of the techniques used by teachers to facilitate children's learning. The approach is based upon a central unifying theme which is referred to as a 'theory of learning,' but this has little to do with what we normally think of as 'learning theory.' It is based upon four fundamental requirements of any instance of learning which are deduced in the course of isolating an operational definition of learning itself. These 'requirements' resemble in part some of the concepts which have been used traditionally, but because they are somewhat different, Mr. Thyne has been courageous enough to coin new terms for them.

The first of these is the requirement of 'cue' which is the essential feature in the learning situation which ultimately initiates the learned response. In terms of conditioning, the cue approximates to the conditional stimulus but, whereas the latter is thought of as evoking a response, the cue is neutral until it is endowed with the second requirement called by Thyne that of 'force.' The satisfaction of the requirement of force determines that the individual will make some sort of response, a third requirement being necessary before it will take a specific form. This is the 'pilot cue' which is akin to the unconditional stimulus of classical conditioning, though again, in itself it does not elicit the response—this is a function of the requirement of force. Like the U.C.S., the pilot cue becomes unnecessary when learning is complete, i.e., when under the fourth requirement, that of 'tie,' the specified, learned form of response is inevitably evoked by the cue.

In a brief paragraph this may sound confused, eccentric and unnecessary; as the ideas are developed in the 60 pages making Part I of the book, the system is seen as coherent, reasonable and practically relevant to classroom situations. The rest of the book is devoted to the application of the system to the job of teaching, which is seen in various roles; as habit-training and breaking, as explaining, as teaching for transfer and for recall, and as training in skills. More traditional aspects are examined in the light of the central theme of the four requirements and general strategies are deduced for classroom practice.



There is no doubt that students who have worked through this book under guidance will have acquired a means of making their teaching more effective. The book is a coherent whole and the parts are logically inter-dependent and lucidly presented. The writing is clear and seems to owe something to programmed learning with its detailed planning, small progressions and frequent summaries. Students should not find the text difficult, but it will demand of them a deal of disciplined thinking. There are a lot of rather lengthy footnotes which seem to have been directed at the lecturer or more advanced student and should be ignored by the beginner, for they deal with more difficult points such as might be raised by an experienced psychologist critically reading the text.

The book will probably gain a lot of adherents; there will be others who will dislike it because they are wedded to a different conceptual framework. It will appeal most to those who, in Wiseman's phrase, wish to make psychology teaching in colleges of education less tender-minded, for as Thyne writes "... a woolly belief in the value of 'experience,' or reiteration of faith in the 'personality' of the teacher, is no substitute for professional competence that can be won only by honest thinking and action based on insight."

L. B. BIRCH.

TOWNSEND, E. A., and BURKE, P. J. (1963). *Learning for Teachers*. Macmillan, pp. 313, no price given.

This is a comprehensive book, dealing with all the usual subjects (motivation, readiness, aspiration, remembering, discrimination, transfer, reinforcement, communication, measurement, etc.). Its aim is "to present the established facts about human learning to those who have chosen teaching as a profession" and it tries "to bridge the gap between the general principles of psychology and courses in methods of teaching."

A great deal of experimental evidence has been sifted as a basis for the text, and restraint has been shown in limiting assertions to what is justified by the evidence. For instance, in the chapter on remembering, we read: "Favourable attitudes predispose persons slightly toward retention. This is not the broad, strong statement of the influence of attitudes which most teachers would like, but it is as much as the facts warrant." Such a quality enhances the general trustworthiness of the statements to be found throughout the book.

As a complete work, however, the book does not seem to the reviewer to succeed. It is jerky and tiresome to read, progressing rapidly over innumerable topics and seldom developing enough smoothness and continuity to sustain interest. In fact, much of it reads like a succession of quotations from summaries of articles. For example, to take a few sentences almost at random: "The learner acquires knowledge and skill with and from people. He performs simple tasks more rapidly when spectators are present than when he is alone (138). The presence of a co-working but non-competitive group increases the quantity of work but not its quality (10). The quality of group results does not surpass the individual results of every member of the group, and generally the top individual comes out better than the group as a whole (258)..." Occasionally, too, the meaning is not quite clear: "An important distinction needs to be made here between maintaining control and dominating. Every person in control dominates, but he may do so for the common good or merely because he enjoys wielding power. This is the difference between being dominant and domineering."

A further factor contributing to the disintegrated effect is the authors' almost aggressively eclectic policy: "There is nothing to be gained by attempting to integrate the principles into a system... When the basic factors are considered as a whole, they are often somewhat contradictory." This may make for unbiased reporting of discreet empirical findings but, unfortunately, it also renders the exposition intellectually less challenging and absorbing.



Moreover, the attempt to form a bridge between teaching methods and pure psychology is not impressive. Too often it takes the form of interspersing commonplace remarks about the classroom or every day life amongst summarized conclusions from experiments.

As has been said, most of the statements made are cautious, but here and there one comes across views, presented as facts, which are at least debatable. For example: "Each disciplinary episode is a specific occurrence, and we help the offender if we can treat it as such. The failure of one person to accede to the control of another has a beginning and an end . . . The preferable situation is one in which each episode is treated as an event; once the penalty is paid, the event is over." Or this: "Children and adolescents transfer their reactions from one controlling adult to another quite indiscriminately. When they learn to individualize and discriminate these reactions, human relationships improve. This is a quality of maturity not to be expected too soon in the young."

All in all, this seems a good enough book to refer to, but not to learn from.

STANLEY NISBET.

## PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

- HEINRICH MENG AND ERNST L. FREUD (Edits.) (1963). *Psycho-Analysis and Faith: The Letters of Sigmund Freud and Oskar Pfister*. Hogarth Press and the Institute of Psycho-Analysis, pp. 152, 21s.
- JAMES M. THYNE (1963). *The Psychology of Learning and Techniques of Teaching*. University of London Press, Ltd., pp. 240, 18s.
- PAUL FRAISSE AND JEAN PIAGET (1963). *Traité de Psychologie Expérimentale. VII—L'Intelligence*. Presses Universitaires de France, pp. 207.
- JOSEPH LUFT (1963). *Group Processes: An Introduction to Group Dynamics*. The National Press, Palo Alto, California, pp. v+57, \$1.95.
- CHESHIRE EDUCATION COMMITTEE (1963). *The Education of Dull Children at the Secondary Stage*. University of London Press, Ltd., pp. 224, 18s.
- C. M. DRILLIEN (1963). *The Growth and Development of the Prematurely Born Infant*. E. and S. Livingstone, Ltd., pp. xv+376, 50s.
- G. W. BASSETT, A. R. CRANE, and W. G. WALKER (1963). *Headmasters for Better Schools*. University of Queensland Press, pp. vii+107, 27s. 6d.
- EDWARD ARTHUR TOWNSEND AND PAUL J. BURKE (1963). *Learning for Teachers*. The Macmillan Co., New York, pp. vi+313.
- F. J. FISH (1963). *Clinical Psychiatry for the Layman*. John Wright and Sons, Ltd., pp. vii+59, 8s. 6d.
- CALVIN W. TAYLOR and FRANK BARRON (Edits.) (1963). *Scientific Creativity: Its Recognition and Development*. John Wiley and Sons, Inc., pp. xxiv+419.
- ARTHUR W. STAATS AND CAROLYN K. STAATS (1963). *Complex Human Behavior*. Holt, Rinehart and Winston, Inc., New York, pp. xiii+546, \$7.50.
- KORNEI CHUKOVSKY (1964). *From Two to Five*. University of California Press, pp. xvii+168, \$4.50.
- JONATHAN BENNETT (1964). *Rationality*. Routledge and Kegan Paul, Ltd., pp. vii+123, 14s.
- Z. P. DIENES (1964). *The Power of Mathematics*. Hutchinson Educational, Ltd., pp. 176, 21s.
- Z. P. DIENES (1964). *An Experimental Study of Mathematics Learning*. Hutchinson Educational, Ltd., pp. xx+206, 50s.
- G. E. PERREN (1963). *Linguistic Problems of Overseas Students in Britain*. British Council English-Teaching Information Centre, pp. 15.
- BRITISH COUNCIL FOR REHABILITATION OF THE DISABLED (Report of a Working Party) (1963). *The Handicapped School-leaver*, pp. 170, 15s.
- DAVID STAFFORD-CLARK (1964). *Psychiatry for Students*. George Allen and Unwin, Ltd., pp. 277, 35s.
- COUNCIL FOR CULTURAL CO-OPERATION OF THE COUNCIL OF EUROPE (1963). *Youth and Development Aid*, pp. 114, 10s.
- ROBERT E. VALETT (1963). *The Practice of School Psychology*. John Wiley and Sons, Ltd., pp. x+339, 57s.
- RAYMOND J. CORSINI AND DANIEL D. HOWARD (Edits) (1964). *Critical Incidents in Teaching*. Prentice-Hall International, pp. xxxii+222, 34s.
- KELLOGG SMITH AND LEIGHTON STEELE (1964). *Proper Punctuation*. English Universities Press, pp. viii+280, 21s.
- CARL F. HEREFORD (1963). *Changing Parental Attitudes Through Group Discussion*. University of Texas Press, pp. xvi+198, \$5.00.
- P. E. VERNON (1964). *Personality Assessment*. Methuen and Co., Ltd., pp. 331, 42s.
- THEODORE LIDZ (1963). *The Family and Human Adaptation*. Hogarth Press, Ltd., pp. 120, 25s.
- LEOPOLD ROSENMAYR (1963). *Family Ties, Social Relations and Leisure Habits of Young Workers in Austria*. (Verlag für Geschichte und Politik, pp. 431, no price.



- R. I. WATSON (1964). *The Clinical Method in Psychology*. John Wiley and Sons, Ltd., pp. xii + 779, 19s.
- PER RAND (1963). *Distortion and Selectivity*. Norwegian Universities Press, pp. xii + 272, N.kr.35.00.
- R. RAY BATTIN AND C. OLAF HAUG (1964). *Speech and Language Delay*. Charles C. Thomas, Illinois, pp. xi + 71, \$4.50.
- JOHN COHEN (Edit.) (1964). *Readings in Psychology*. George Allen and Unwin, Ltd., pp. 414, 52s. 6d.
- MARCELLE KELLERMANN (1964). *Two Experiments on Language Teaching in Primary Schools in Leeds*. Nuffield Foundation, pp. vii + 77.
- R. S. MACARTHUR, S. H. IRVINE AND A. R. BRIMBLE (1964). *The Northern Rhodesia Mental Ability Survey, 1963*. Rhodes-Livingstone Institute, Lusaka, pp. viii + 72, 21s.
- PAUL R. FARNSWORTH, OLGA McNEMAR, QUINN McNEMAR (Edits.) (1964). *Annual Review of Psychology*. Annual Reviews, Inc., Palo Alto, California, pp. vii + 629, \$9.00.

## CORRECTION

### PROMPTING AND FEEDBACK IN THE LEARNING, RETENTION, AND TRANSFER OF CONCEPTS

BY M. C. WITTROCK and P. A. TWELKER  
(University of California, Los Angeles)

- (1) Page 10 : Summary : line 6 should read " As hypothesized, KCR increased learning, retention, and transfer only when the rule was not given."
- (2) Page 16, 3rd para., line 5, should read : " The KCR variable produced no dependable effect on affectivity ratings, . . . "
- (3) Page 16, Table 5, third column : " Retention and Transfer Test " refers to " Time " and not to " Attitude."

# Programmed Learning

*Journal of The Association for Programmed Learning*

Editor : **JOHN ANNETT**

*Psychology Department, University of Aberdeen.*

The aim of the Association for Programmed Learning is to promote the investigation and use of the new teaching techniques which have come to be known generally as Programmed Learning. The first issue of this journal was published in May 1964.

These new teaching techniques are developing rapidly in this country and are going to be of ever increasing importance in all fields of education. They are beginning to be widely used by teachers, local authorities, industrialists and the armed forces. An understanding of them and an interest in their future development is therefore of particular importance for educationalists, psychologists and scientists.

*Programmed Learning* is a journal which will, in effect, provide a forum where education, psychology and science can meet and discuss problems of common interest. It will contain : original theoretical articles, research reports, notes on special devices and techniques or unusual applications, abstracts of articles published elsewhere, programmed evaluation data, survey of current literature and book reviews and the official proceedings of the Association.

*The first issue contained the following articles :*

A HISTORY OF A PROGRAM ON PAPERMAKING by B. T. Dodd and R. Mitchell.

THE EFFECTS OF LINEAR AND BRANCHING METHODS OF PROGRAMMED INSTRUCTION ON LEARNING AND RETENTION OF A TOPIC IN ELEMENTARY SCIENCE by T. C. Larkin and G. O. M. Leith.

THE AUTO-TUTOR AND CLASSROOM INSTRUCTION by Peter Cavanagh.

THE FEEDBACK CLASSROOM by K. Holling.

PROGRAMMED LEARNING IN EMERGING NATIONS by J. Hartley.

THE ROYAL NAVY STUDY by D. Wallis and R. P. Wicks.

Three issues a year : May, July, November

Annual subscription 21s. Separate parts 7s. 6d.

10s. 6d. to members of the Association for Programmed Learning.

Copy of the first part available on approval.

## Sweet & Maxwell

11, New Fetter Lane, London, E.C.4.







CHARLES WILFRID VALENTINE (1879-1964)



## CHARLES WILFRID VALENTINE (1879-1964)

From the stop-press slip in the preceding issue readers will already have learned to their deep regret that Professor Valentine, who founded this *Journal* and edited it for twenty-five years, died at his home near Birmingham, on May 26th. By his death, both British psychology and British education have suffered a heavy loss. Few names, I imagine, have become so well known to teachers and educationists during the past quarter of a century, or have been so widely cherished and respected.

Charles Wilfrid Valentine was one of eight children. His father, the Rev. Henry Valentine, was a Methodist Minister; and in accordance with the circuit principle, the family was continually on the move. Charles was born at Run-corn; started education at Wakefield, continued it at Nottingham and Preston Grammar Schools, and at seventeen went direct from school to teaching in a small boarding school. Thanks to his savings he was able to enrol as a student at Aberystwyth, where he read for the London B.A. After a further spell of teaching, he entered Downing College, Cambridge, as a Foundation Scholar, and eventually gained a double first in the Moral Science Tripos.

On the advice of Professor Ward, under whom he had studied, he went to Germany "to learn something at first hand of the new experimental methods"; and it was in the psychological laboratory at Würzburg that I first met him, and a lifelong friendship was formed. The students who came to work under Külpe were a cosmopolitan crowd. There were Germans (nearly all older than ourselves), a few Austrians, one Italian, two Americans and four Englishmen, Valentine, Flugel and myself, and a little later, Pear. H. J. Watt (who followed Woodworth and preceded me as Sherrington's assistant at Liverpool) had just left after publishing an epoch-making paper on "the experimental study of thought-processes," which, with equally notable papers from Marbe and Ach, had initiated what figures in the histories as "the Würzburg school": their work, as Valentine himself has observed, "marked the first break-through of experimental psychology into the new and higher realms of thinking and problem solving." Another famous student had also just left after taking his doctorate—Max Wertheimer, who founded the still more famous "school of Gestalt." Rather to our surprise, however, we found that Külpe himself was now far more interested in the psychology of aesthetics. Indeed, the prevailing trend of our contemporaries there was to move away from the strictly experimental approach of the Wundtian laboratory, and rely more and more on methods of "trained and planned introspection"—methods which had already extended the unsuspected notion of the contents of consciousness and demonstrated the cognitive as well as importance of unconscious attitudes and processes in the cognitive as well as the motivational field.

Valentine's gay humour and high spirit introduced a new and unfamiliar element into the life of the laboratory, and did much to smooth away the language difficulties. Professor Külpe, however, always thought it necessary when introducing a newcomer, to explain that "Der Englischer Herr—he is such a Great Komik." In separate cubicles Valentine and I were first set to carry out experiments on memory with the Ebbinghaus technique—repeating over and over again a list of twelve or more nonsense syllables presented in monotonous succession by a carefully timed machine. The story goes that Külpe, looking in every quarter of an hour, finally expressed surprise that, although half the morning had passed, he still had not learnt them. "Oh!", cried Valentine, with a gasp of relief, "I didn't understand I was meant to learn them." And, in



five more minutes he had them by heart. The incident was freely quoted as an unexpected confirmation of Watt's emphasis on the *Aufgabe* (the 'motivating task'); and Külpe recommended us to read the little book by his own fellow-student, Ernst Meumann, on 'Economy and Technique of Learning.' This, in turn, led us later both to purchase Meumann's two volumes on *Experimentelle Pädagogik*, which appeared during our stay; and we spent many happy evenings in the neighbouring *Biergarten* discussing how far a new science of experimental pedagogy was really practicable—discussions in which Külpe himself frequently joined.

On his return to England, Valentine was offered a post as Demonstrator in Myers' laboratory at Cambridge; but he ultimately decided to accept a lectureship at St. Andrews. Here, largely with the theories of Lipps and Külpe in mind, he embarked on a long series of investigations on aesthetic appreciation in both the visual and the auditory fields: it was a subject (or, as he sometimes modestly put it, a "hobby") which occupied him right up to the end of his life. At Stout's suggestion, the results of these early experiments were brought together in Valentine's first book, *An Introduction to the Experimental Study of Beauty* (1913)—a small cloth-bound volume which then cost only sixpence.

There is no need to enumerate his many publications: a full list has already been printed in this *Journal* (XXVI, pp. 3-12), together with a detailed appraisal of contributions. They comprise more than fifty papers and sixteen books, and cover an astonishingly wide range of topics—mental tests, the function of imagery, the transfer of training, the nature of temperament, the reliability of examinations and of character-judgments, discipline at home and in the classroom, and the educational value of Latin and Greek—nearly all of them issues which, as topics for scientific investigation, were relatively new when he first took them up. Nor were his interests confined solely to the problems of education. The catalogue includes highly original studies of such diverse subjects as the religious philosophy of Lotze, various optical illusions, dreams and the unconscious, the psychology of laughter, and (among his war-time investigations) the principles of army instruction.

Of all his many publications, probably the most influential has been the comprehensive volume entitled *Psychology and its Bearing on Education* (1950). It at once became a standard textbook for students in training colleges. It was reprinted eight times, and then, revised and brought up to date, reappeared as a fresh edition in 1960. It is, I understand, to be re-set as a paper-back, largely with an eye to the requirements of students overseas. Meanwhile, his book on intelligence tests has been translated into both Italian and Greek. His last book, *The Experimental Study of Beauty*, published when he was 83, had virtually the same title as his first, and was extended to nearly 450 pages. It has already been hailed as a classic in this field.

He was as versatile in his use of methods as in his choice of topics. Unlike most psychologists, both then and now, he placed his faith in no one type of procedure, but was thoroughly eclectic. From Myers and Rivers, he had learned the value of experimental techniques; from Külpe and his colleagues at Würzburg the value of systematic introspection; and he was one of the first to apply the new statistical devices developed by Pearson in London and Yule at Cambridge. But the procedure in which he particularly showed his ingenuity and skill was that of planned observation, a procedure which is so effectively employed in his work on *The Psychology of Early Childhood*, and which deserves to be copied more widely by investigators of the present day. In some of his most important studies he uses all four methods in combination.



His name, however, for the vast majority of educationists and teachers is, and always will be, linked with that of the *British Journal of Educational Psychology*. This, as more than one overseas psychologist has told us, rapidly became the leading journal in its own special field. Just over fifty years ago, with the help of what was then the Training College Association, Professor J. A. Green of Sheffield started a *Journal of Experimental Pedagogy*, modelled on the *Zeitschrift für Experimentelle Pädagogik* which Meumann had successfully launched six years before. When Green died shortly after the War, Valentine was invited to assume the editorship; he readily assented, provided the somewhat forbidding title was changed. The journal was accordingly re-christened *Forum of Education*, and its scope correspondingly enlarged. During the years that followed, several new periodicals appeared, dealing with the more general issues—administration, history and philosophy of education, teaching methods, and the like. In 1931, therefore, the British Psychological Society approached the Training College Association with the proposal that the *Forum* should be replaced by a larger journal, sponsored by both bodies, and devoted exclusively to educational psychology. This was agreed. But the two societies were informed by their legal advisers that the surest way to achieve their purpose would be to form an independent limited liability company, which would act as the proprietors and publishers of the new venture. As a result, certain psychologists and educationists suddenly found themselves playing the unfamiliar role of company promoters. There were six shareholders, a board of six directors, and a capital of £100. As soon as the necessary formalities had been fulfilled, the directors promptly selected a managing director who, in the interests of neutrality, was *not* a shareholder, and appointed him editor. Their choice, needless to say, fell on Professor Valentine. The remarkable success of this curiously circuitous procedure, coupled with the policy of sticking to a good editor once you have found him, is a device that seems to merit the consideration of other bodies who may wish to establish a new scientific journal in these difficult times.

That the journal so quickly reached financial stability was due entirely to the devotion and energy of Valentine himself, who, it should be remembered, was at the same time running a large education department with the aid of a very small staff, and composing articles and books of his own. What it meant to be publisher and managing director as well as editor could best be appreciated by visitors who happened to call at his home when a new number was about to appear. I have witnessed fifteen hundred envelopes or more being addressed by Mrs. Valentine; stacks of copies fresh from the printers, waiting on the sideboard, to be inserted by one of the children; and Professor Valentine himself rushing off to the nearest post office with the first of a well-filled series of sacks. It was only when the circulation ultimately reached two-thousand that this mechanical drudgery was at length transferred to the officers of Messrs. Methuen and Co., who have acted as publishers since 1948.

In Valentine's hands the journal became, in standpoint as well as in name, pre-eminently a *British* journal. "The problems of British schools and of British education," so he reminded potential contributors, "are characteristically British problems; and it is the job of British investigators to seek the solutions—they cannot be imported from abroad." And here he spoke from an experience of such problems gained at first hand as teacher in four or five different types of school in widely different parts of the country. With McDougall he continually deplored what McDougall was wont to call "our unjustified insular modesty." In his Presidential Address to the Society, he

vigorously criticized the growing tendency of British writers to lay major emphasis on the teaching of this or that "school of psychology"—nearly all of them, as he says, having their origin abroad, according as the fashion favours some French, German, Austrian, Swiss or American leader. "Even when some pregnant idea, put forward by some early British psychologist, is endorsed and advocated, it is nearly always presented as emanating from some other country and is given its foreign label." And he pleads for "an authoritative volume by a representative team on psychology *as expounded in this country*." (We are almost the only leading country that lacks such a volume!) A year or two later he himself set the example in the textbook already quoted: there 80 per cent. of the many researches and articles referred to are British, whereas in nearly all other current textbooks, the proportion of British to foreign work is reversed.

Above all, he was, as his present successor has expressed it, "essentially a creative editor." He was never content to wait for contributions to be submitted, but was himself continually suggesting topical problems for discussion, often in the form of Symposia, and still more frequently soliciting the help of eminent authorities to write on the outstanding issues of the day. In this way he undoubtedly exercised a powerful influence on educational policy and opinion, and did much to direct the course of educational research.

But, wide as has been the sway which he wielded as research worker, author and editor, his influence, I believe, has chiefly made itself felt through numerous personal contacts. Whether in private or on the platform, whether dealing with students, colleagues, teachers, or educational officials, he at once impressed all who saw or listened to him with his vivid and forthright personality. You were struck immediately with his willingness to help, his sympathetic understanding, his quick and ready humour, and, from first to last, his intellectual honesty. And behind all this there lurked a shrewd and penetrating mind, and a sound, business-like, and knowledgeable judgment. Above all, as one of his own students puts it—and as the children he talked to or tested seemed at once to realize—"he was so very loveable."

CYRIL BURT.



## AN ITEM ANALYSIS OF THE ADVANCED MATRICES

BY A. R. FORBES\*

*(Department of Psychological Research, The Crichton Royal, Dumfries)*

SUMMARY. Prior to revising Advanced Matrices (1947) an Item Analysis was carried out on a sample of cases drawn from a parent population of over 2,000. The method of analysis is described, and the reasons for the retention and rejection of items given. The findings are discussed in the light of those of Yates (1961). It is concluded that the new edition, Progressive Matrices (1962), although shorter will be at least as efficient as the 1947 version.

## I.—INTRODUCTION.

FOULDS AND RAVEN (1950) reported the consistency, re-test reliability and psychological validity of the Advanced Matrices (1947) as a forty-minute test given in all to 1,900 people. From different scholastic and cultural levels, they selected groups of children over 10 years of age, and young people in various training colleges and university faculties. They found that, as a test of intellectual efficiency, from the age of 11 onwards, it had high retest reliability and that, unlike the Standard Progressive Matrices (1938), for university students, it produced no 'ceiling effect' at the end of the Scale or at any intermediary point.

Yates (1961) reported an Item Analysis of the Advanced Matrices (1947) scores of 187 university students divided into two matched groups. One group (N=61) were allowed thirty minutes working time, the other group (N=126) forty minutes, in which to do the test. He found that in both groups the first 20 items were correctly solved by almost every student. From item 36 onwards there was a sharp fall in the pass rate of attempted items, coupled with an increase in the number of items not attempted. Items 20-36 he found highly discriminative, although items 21-24, and again 33-36 appeared to be mis-placed in order of difficulty. Yates concluded that for his university student populations at least, items 1-20 could be dropped and the remainder re-arranged as a 'power test.'

The present item analysis was designed:

- (1) To check the behaviour of the individual items constituting the 1947 series when the test is given with a forty-minute time limit to a large and representative sample of young people of more than average intellectual ability.
- (2) To find which items in the Scale could be eliminated without detriment to the test as a whole.
- (3) To determine how the remaining problems varied in order of difficulty with a view to preparing a new edition of the Scale.
- (4) To examine the erroneous figures chosen to complete each problem and to note any differences in the nature of the errors made at different score levels.

It was also designed to show how far Yates' findings, based on a highly selected sample of university students and using two time limits, are generally applicable. It must, however, be borne in mind that according to Raven (1948), Advanced Matrices (1947) was developed from the Standard Progressive Matrices (1938) for use with everyone of more than average intellectual ability as a

\* Now at the Medical School, University of Otago, N. Zealand.

non-verbal test which could be used without a time limit in order to assess a person's total capacity for coherent perception and orderly thinking, or with a time limit in order to assess his intellectual efficiency when working under stress, and to this extent to overcome the unavoidable limitations of the Standard Scale.

## II.—THE POPULATION STUDIED.

The parent population for this study consisted of 2,256 protocols obtained from three groups of testees. The first group of 1,500 cases, consisted of men passing through the R.A.F. Aircrew Selection Centre; the second, 500 in all, were youths tested in pursuance of a selection procedure enquiry by the G.P.O.; the remaining group, containing men and women, were students at a teachers' training college. The age distribution for the total population was trimodal. The G.P.O. group was the youngest, being telephone engineering applicants 15-16 years of age. The R.A.F. sample consisted almost entirely of men aged between 17 and 20. The age distribution for the teacher-training college sample was positively skewed, with the median in the early 20's, but with a few cases over 30 years of age. For the purpose of the analysis, it was considered that these three groups covered the whole age range for which the test was designed.

From this parent population, samples of up to twenty-five test protocols were drawn at every possible total test score between 0 and 48. For the item analysis envisaged, samples of this size were considered sufficiently large for statistical treatment without giving undue weight to performances around the mean for the population as a whole. For each total score between 21 and 41, inclusive, it was possible to obtain a sample of this size. One hundred and ten people obtained scores of less than 21. These were distributed as follows: Nineteen people obtained a score of 20, twelve people a score of 19, twelve people a score of 18, eleven people a score of 17, eight people a score of 16, and eleven scored only 15. In addition to this there was a small but decided cluster around a total score of 7, leading one to suspect that these were people who had found their way into the parent populations for reasons other than that of intellectual ability. This is supported by an earlier finding (Raven 1941). At the upper end of the Scale only eighteen people achieved a score of 42, six a score of 43, three a score of 44, two a score of 45, and one a score of 47. As it was important to know as much as possible about the behaviour of the test at the extremes, the detailed item analysis was extended to cover all available score groups between 16 and 42, inclusive.

Within this score range, protocols were drawn from the parent population in the following way: from the G.P.O. and training college protocols in the ratio as far as possible of their total sizes, i.e., 2:1. Where this was not sufficient, the samples were made up with additional cases from the R.A.F. population which was used as a reserve.

## III.—METHOD OF ANALYSIS.

An item analysis which consists of simply correlating performance on each item with a total score on the scale as a whole fails to show how, for example, people of different levels of ability cope with the test. It also tells us nothing about the distribution of choices amongst the eight available figures from which selection has to be made. In view of this, it was agreed that the method of item analysis to be used in this study should conform to that used for earlier analyses of the Matrices scales. This requires that each group of subjects obtaining a given score should be treated as a representative sample of persons



of that level of 'Matrices ability,' and that the behaviour of each item should be examined at each such score-level. Only on the basis of this detailed analysis of item behaviour at every scoring level can a comprehensive picture of item behaviour over all the range of scores be synthesised.

So that no information should be lost, and so that the data could be presented in the most convenient form for comparison with earlier item analyses, work sheets were prepared as follows:

Large sheets of quarter-inch squared paper were divided into forty-eight columns, one for each of the forty-eight test items constituting the scale. Across each of these sheets lines were ruled at intervals of eight squares, one for each of the eight figures between which choice had to be made in each problem attempted. Thirty rows each of forty-eight cells formed in this way were prepared, i.e., one for each total score 16-45 for which data were available. By making the appropriate entry in each of the cells formed in this way it was possible to record for each of the forty-eight problems in the test every right or wrong figure chosen by every person in the sample studied.

The information obtained in this way made it possible to read off at any given score level:

- (a) The number of people who chose the correct figure to solve each problem.
- (b) The frequency with which each of the other seven erroneous figures was chosen.
- (c) The frequency with which any given problem was not attempted.

From the same work sheets it was also possible to determine:

- (d) Changes in the frequency with which any given problem was solved as the total score on the scale increased.
- (e) Any significant changes in the types of erroneous figures chosen.

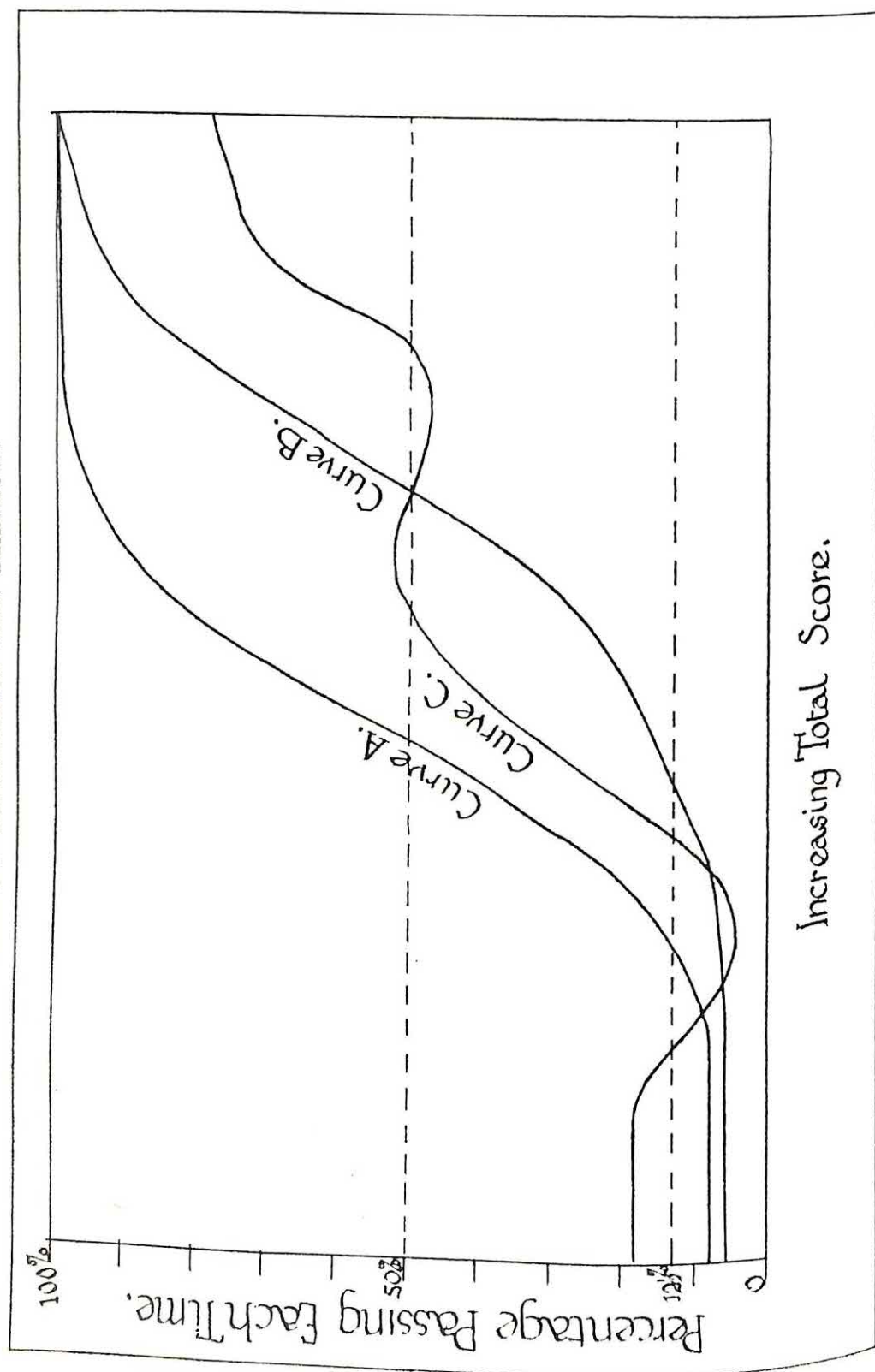
By summing the totals in each row and column, it was also possible to check errors of tabulation, and finally to prepare graphs showing for each item in the scale the rate of change in the percentage of passes as the total score on the scale increased from 16 to 42, inclusive.

#### IV.—RESULTS.

##### (1) *Item Behaviour.*

If each item constituting the test assesses the same intellectual functions, but at progressive levels of ability, Graph 1 shows how, in principle, efficient or inefficient items would behave as the total score on the test increased from zero to its maximum. At any given point the rate of change in the percentage of passes on any item shows the correlation of passes on that item with increase in total score, while alterations in the rate of percentage of passes as the total score increases will indicate defects in the problem or in the alternative figures between which choice has to be made. In Graph 1, curve A represents the ideal behaviour of a psychologically efficient test item of less than average difficulty; curve B the behaviour of an equally efficient test item of greater difficulty; curve C represents the behaviour of an inefficient test item in the same area of difficulty as curve B. The broken line at the  $12\frac{1}{2}$  per cent. pass level represents the frequency with which the correct figure could be chosen at random out of a group of eight figures.

GRAPH 1—HYPOTHETICAL ITEM BEHAVIOUR





Curve A, representing the behaviour of an efficient test item, stays in the area of random scoring until it begins to be psychologically problematic. After this it rises very steeply and uniformly as the total score on the scale increases. Throughout the 40-60 per cent. pass band its rate of change with total score remains constant. Beyond the 80-90 per cent. pass level the rate of change resembles that at which it emerged from the 10-20 per cent. pass level. After it reaches the 100 per cent. pass level, no further failures in it occur. The sharp rise in the middle range implies that there is some important aspect in the process of solving this problem, which only emerges in that level of problem solving ability as measured by the total test score. Before this level of ability is reached, practically no one chooses the right figure. After it has ceased to be problematic, no one is careless enough to select a wrong figure to complete it. Curve B, representing the behaviour of an equally efficient but more difficult test item, shows the same sharp rise from below random scoring to 100 per cent. passes, but at a significantly higher level of problem solving as measured by the same scale.

Curve C representing the behaviour of an inefficient test item shows the four most serious defects of problem design. Before this test becomes psychologically problematic, the right answer is selected far more often than is expected on a random basis. In other words, the right figure attracts a person's attention more than the other seven figures shown with it, either because of its position or because of its design. For some reason or other, just before this test item begins to be problematic, the percentage of successful guesses falls to well below the  $12\frac{1}{2}$  per cent. to be expected by random selection. This may occur if at this stage one of the erroneous figures becomes a more attractive choice. After this test item begins to be psychologically problematic, the percentage of passes only rises to around the 50 per cent. pass level. For a considerable range of ability it fluctuates about the 50 per cent. pass level until a stage is reached at which 60 per cent. of testees solve the problem correctly. The erroneous choices made at this level of ability may show that in this area people are really choosing one out of two closely similar figures. Beyond the 60 per cent. pass level curve C still rises slowly and never attains a 100 per cent. pass level. In other words, there are always people who do not trouble to be sure exactly which out of the two more or less similar figures is quite correct. It should be noted in passing that the area of the Graph below curve C is almost identical with the area below curve B. It follows that from the total percentage of passes, these two problems would appear to be of equal difficulty. Only by using a large sample and a wide score range does the difference between them become clear.

Two kinds of obviously inefficient test items are not shown in Graph 1. These are problems which are either solved by everyone without exception or those which are solved by virtually no one given the test.

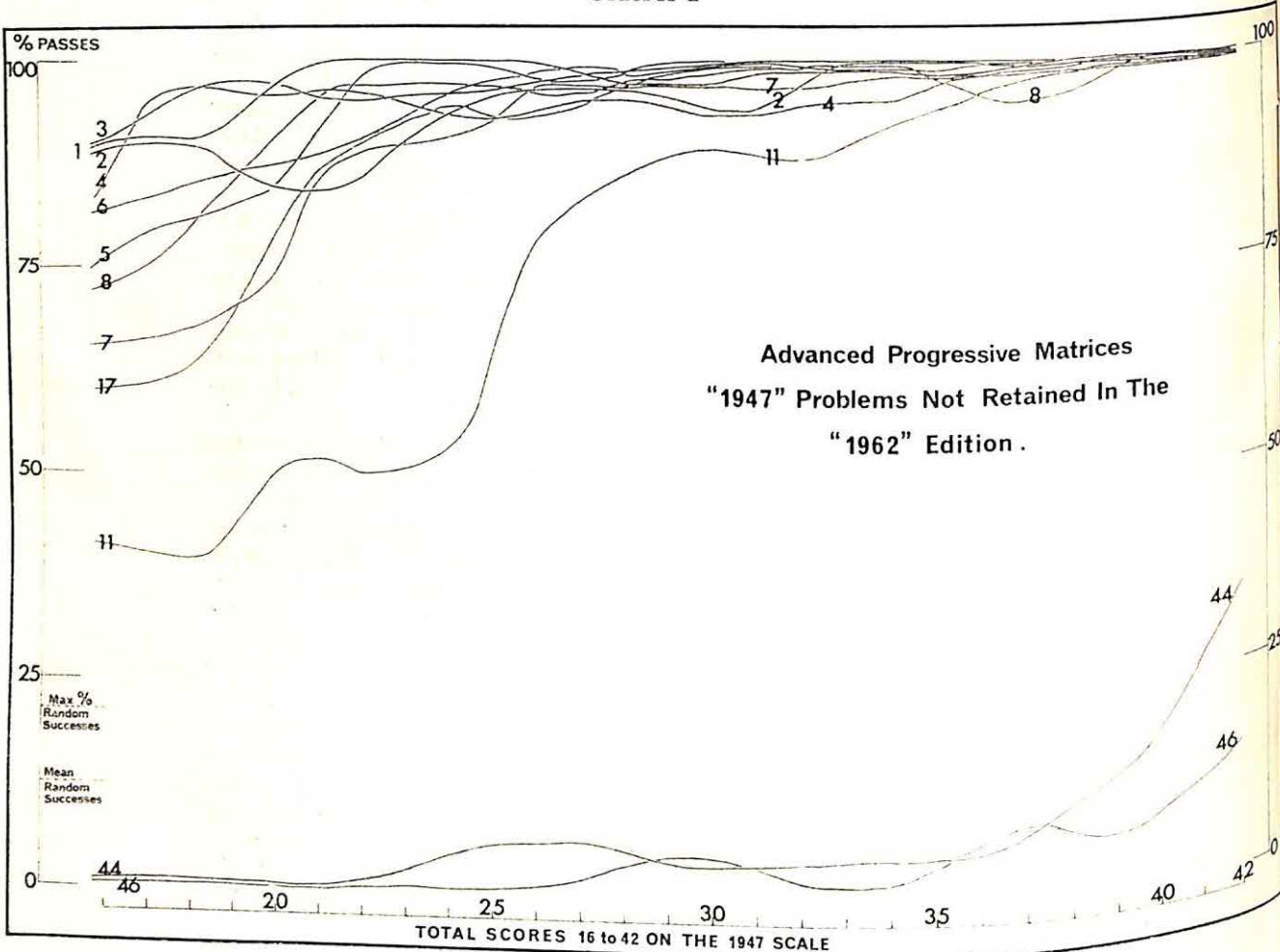
## (2) *Items to Eliminate.*

Graph 2 shows the behaviour of twelve items in the 1947 edition which could be discarded without detriment to the Scale as a whole. Numbers 1 to 8 were passed by practically everyone scoring more than 18 out of the total of 48. As Yates found, these make virtually no contribution to the total score.

Item 11 shows all the characteristics of an inefficient test. For the population studied, it was always solved by more than 40 per cent. of the testees. Between the score of 18 and 25, it fluctuated around the 50 per cent. pass level. In addition to this, at a total score of 30 or more, 12 per cent. of the testees were

GRAPH 2

Advanced Progressive Matrices  
 "1947" Problems Not Retained In The  
 "1962" Edition.



still making wrong judgments. Problem 46 was almost never solved by anyone scoring less than 42 out of 48 (approximately 5 people out of 2,256). Problem 44, although difficult, seemed to behave well. Unfortunately, the argument presented was found to be indeterminate and had, for this reason, to be removed from the scale.

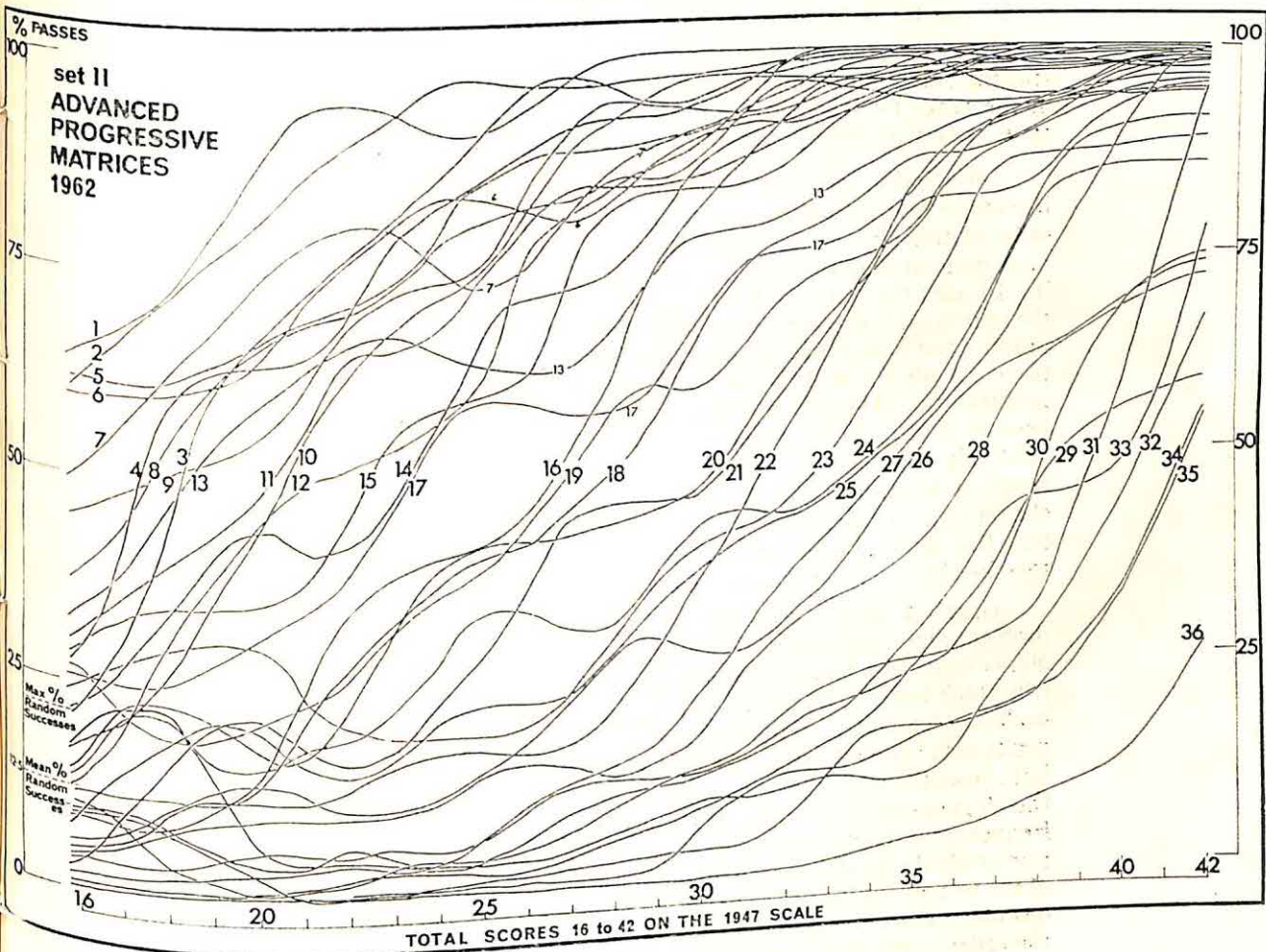
### (3) Items to Retain.

Graph 3 shows for each of the remaining problems the increase in percentage of passes as the total score on the 1947 edition increased from 16 to 42 inclusive. The number inserted on each curve shows the position of the problem in the 1962 series, not its original position. Its position in the 1947 series can, however, be read off from Graph 4 or Table 1.

At successive levels of difficulty problems 3, 10, 16, 27 and 31 of the new series approximate to the idealised item behaviour curve. Problem 3 is passed by less than 20 per cent. of the subjects scoring 16 on the 1947 scale. From this point onwards passes on this test rise rapidly as well as uniformly to the



GRAPH 3



85 per cent. pass level. After this, the form in which the problem is presented encourages a little guessing before it reaches the 100 per cent. pass level. Problem 10 behaves in the same way except that it is more difficult and does not encourage guessing. Problem 16 shows a slight plateau just before it reaches the 40 per cent. pass level. From this stage onwards, it rises rapidly to 95 per cent. passes where again it encourages a little guessing. Problem 27 gives an almost ideal curve. The same is true of problem 31 with the exception that the position of the right answer (position 4) encourages a high percentage of random successes before a rational explanation of the choice can be arrived at. Apart from this it is probably one of the best as well as one of the most difficult problems in the scale.

Of the problems shown in Graph 3 to be retained in the 1962 edition, the curves for items 6, 7, 13 and 17 are the least satisfactory. Those for items 6 and 7 show prolonged plateaux around the 75 per cent. pass level. The curves for problems 13 and 17 show two plateaux, one near the 60 per cent. pass level, the other around the 75 per cent. pass level. Problem 34 had to be re-drawn



in order to eliminate irrelevant data causing prolonged delay in arriving at and confirming the nature of the correct solution.

From Graph 3 it is clear that when we include technical as well as university students, all of these thirty-six problems effectively discriminate between testees in the range for which the Advanced Matrices was designed. No item in this series is ever passed by everyone and no item is always failed.

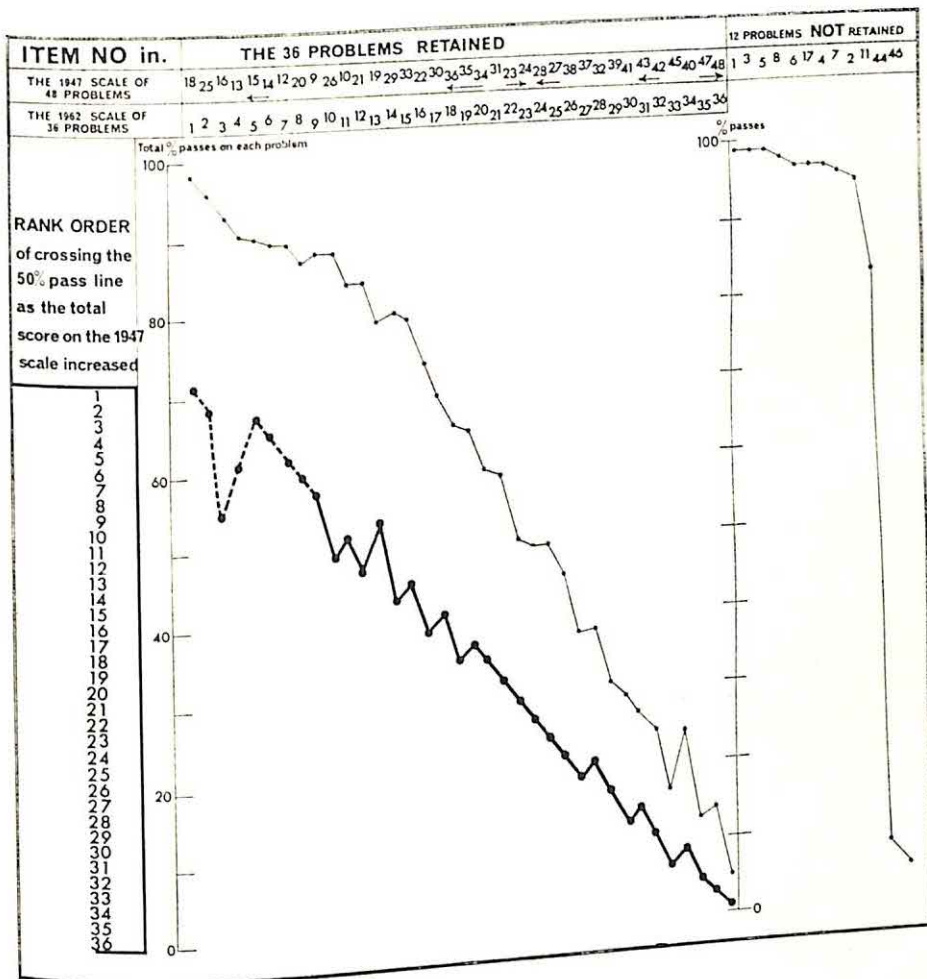
Considering any test item which is good in itself in that it behaves more like curve A or B than curve C in Graph 1, it may still be out of place as regards order of difficulty. A position needs to be found for it in the scale such that it is more difficult than the item preceding it and less difficult than the one succeeding it. Ideally, the criterion for this would be such that there is no overlapping of the curves on the graph. It is impossible to achieve this degree of perfection in practice but a useful minimum requirement is that any overlapping should take place outside the 40 to 60 per cent. pass band where the test is psychologically problematic. The new numbers 1 to 36 show the extent to which this has now been achieved. In the new series, apart from problems 1, 2, 5 and 6, which are relatively easy, the remaining problems traverse the 40 to 60 per cent. pass range at relatively uniform score increments. In no case did the problems cluster together according to the nature of the argument presented. In view of this the original idea of keeping qualitatively distinct groups of problems together had to be reconsidered.

Graph 4 shows a comparison between the total per cent. passes for each problem (the thin line), and the order in which each problem reached the 50 per cent. pass level as the total score in the new series increases from 1 to 36 (the thick line). The top line of figures above the Graph shows the number of each problem in the 1947 edition which was retained. The second line shows its position in the new series of thirty-six problems. The column of figures on the left shows the rank order in which each retained problem was found to traverse the 50 per cent. pass level. The scale to the right shows the total percentage frequency with which each problem on the new series was solved by the population studied. Problems 1, 2, 5 and 6 were never solved by less than 60 per cent. of the testees studied. These are, therefore, shown in the Graph by a discontinuous line. It will be seen that after problem 6 there is a close agreement between the order in which the problems traverse the 50 per cent. pass level and the total percentage frequency with which the problems were solved. The curve to the extreme right of this graph shows the total percentage passes on each of the twelve items eliminated, and their serial numbers in the 1947 edition.

Before deciding on the final order in which to arrange the thirty-six problems retained in the new revision, two prototype tests were made up. In one of these the problems were arranged in order according to the percentage frequency with which they were solved. In the second, they were arranged as far as possible in order of argument. Six psychologists were asked to work through each series and record their introspections. It was unanimously agreed that the total percentage frequency with which the problems were solved was the better order of presentation. One or two exceptions had to be made to this ruling. Problem 13 (formerly 19) has already been mentioned as one of the least satisfactory problems in the revised scale. From Graph 4 it was thought to be a little easier than the total percentage passes would suggest. In order to see if practice with problems presenting similar arguments did, in fact, increase the percentage of passes, it was decided to keep the 1947 problems, 41, 43 and 42 (30, 31, 32 in the 1962 edition) together, although it meant displacing the old



GRAPH 4



problem 45 by one position. As the 1947 problem 40 (now 34) had to be redrawn to eliminate irrelevancies, it was decided that it should come before the 1947 problem 47 (now 35).

#### (4) *Common Errors.*

Before revising the 1947 edition of the Advanced Matrices it was important to know the errors made in attempting to solve each problem and how these were distributed between the eight figures from which one had to be selected. This was not undertaken by Foulds and Raven in their original item analysis. As the Advanced Matrices is designed to be essentially a test of clear accurate thinking, if necessary under the stress of time, the main object of comparing the errors was to ensure that from the time any given item becomes 'psychologically problematic,' there should be no misleading type of erroneous figure, whereas before this effective range is reached a person should never choose the right figure for wrong reasons.

In order to study disorders of thinking, the use of the Coloured Progressive Matrices, Sets A, Ab, B is recommended. Raven (1960) showed that when people had a capacity for coherent perception and orderly thinking insufficient to solve a given type of problem, they usually fell back on repeating a figure already shown to them. A person of greater intellectual ability adopted the principle that the most complex figure must be the right one, or adopted some more or less arbitrary method of reasoning. At the level of ability at which a person was just unable to grasp the right principle of reasoning, he frequently chose a figure which was right as far as it went but was, nevertheless, not quite correct.

In the Advanced Matrices, four types of error could, in fact, be identified. (A) Incomplete solutions. These were errors due to people failing to grasp all the variables determining the nature of the correct figure required to complete a test item. Instead, they chose a figure which was right as far as it went but was only partly correct. An example of this is seen in problem 13 in the 1962 revision for which people chose figure 6 instead of figure 2. Keeping to the terminology used by Raven (1960) this type of error is referred to in Table 1 as an *Incomplete Correlate*. (B) Arbitrary lines of reasoning. Here the figure chosen suggests that the person has used a principle of reasoning qualitatively different from that demanded by the problem. An example of this is seen in problem 16 for which people chose figure 1 instead of figure 4. This type of error will be referred to as *Wrong Principle*. (C) Over-determined choices. These were errors involving failure to discriminate irrelevant qualities in the figure chosen, and to choose a figure which combined as many as possible of the individual characters shown in the matrix to be completed. An example of this is seen in problem 7 for which people chose figure 2 instead of figure 6. This used to be classified under the heading of 'incomplete individuation.' It may be regarded as a form of 'over-inclusive thinking,' but as people who are certainly not schizophrenic make errors of this type they will be described as due to *Confluence of Ideas*. (D) Repetitions. These are errors made by people who simply selected a figure identical with one of the three figures in the matrix immediately adjacent to the space to be filled. An example of this is seen in problem 6 for which people chose figure 7 instead of figure 1.

It was noticed that the types of erroneous choices varied as the total score on the test increased. Therefore, in order to compare the errors made at successive levels of intellectual ability, the revised Scale as a whole was divided into three equal parts. For problems 1-12, inclusive, a comparative study of erroneous choices was made at the score level 7-9, inclusive. For problems 13-24, inclusive, it was made at the 17-19 score level. For problems 25-36, it was made at the 27-29 score level. In the 1947 edition these corresponded to the score levels of  $20 \pm 1$ ,  $30 \pm 1$ ,  $40 \pm 1$ .



TABLE 1

POSITION AND NATURE OF COMMONEST ERRORS MADE AT THREE LEVELS OF ABILITY.

A—Incomplete Correlate.

C—Confluence of Ideas.

B—Wrong Principle.

D—Repetition.

Problem No.	Percentage Passes	Correct Figure	Commonest Error		Second Commonest Error	
			Position	Nature	Position	Nature
LEVEL 1.						
1 (18)	98	5	2	A	1	A
2 (25)	96	1	2	D	{ 6	C
3 (16)	93	7	6	C	7	A
4 (13)	91	4	2	B	4	B
5 (15)	90	3	7	A	8	B
6 (14)	89.5	1	7	D	1	A
7 (12)	89.5	6	2	C	2	D
8 (20)	87	1	4	D	4	D
9 (9)	88	8	3	B	5	C
10 (26)	88	4	8	B	4	B
11 (10)	84	5	4	B	7	C
12 (21)	84	6	5	B	1	A
					{ 4	B
					7	B
LEVEL 2.						
13 (19)	79	2	7	A	{ 6	A
14 (29)	80	1	5	B	8	D
15 (16)	79	2	4	C	4	B
16 (22)	73	4	1	B	6	A
17 (30)	69	6	3	A	5	C
18 (36)	65	7	5	A	4	A
19 (35)	64	3	5	A	1	A
20 (34)	59	8	2	A	8	A
21 (31)	58	8	1	A	4	C
22 (23)	50	7	8	C	4	B
23 (24)	49	6	8	A	2	A
24 (28)	49	3	1	B	5	B
					8	B
LEVEL 3.						
25 (27)	45	7	2	B	8	C
26 (38)	37.5	2	6	A	5	A
27 (37)	38	7	4	B	—	—
28 (32)	32	5	4	A	3	A
29 (39)	29	6	2	A	7	A
30 (41)	26	5	4	A	8	A
31 (43)	24	4	8	A	2	A
32 (42)	16.5	8	4	A	5	A
33 (45)	24	5	7	A	4	B
34 (40)	13	1	3	B	2	A
					{ 4	B
35 (47)	14	3	1	B	7	C
					8	B
36 (48)	5	2	1	B		

Table 1 shows the position and nature of the two most common errors made at three score levels on the Advanced Matrices (1962).

TABLE 2

PERCENTAGE FREQUENCY OF TYPES OF COMMON ERRORS AT SUCCESSIVE LEVELS OF ABILITY.

Type		Level			Total
		1	2	3	
Incomplete Correlate	(A)	23.1	52.0	58.4	44.0
Wrong Principle	(B)	38.5	28.0	33.3	33.3
Confluence of Ideas	(C)	19.2	16.0	8.3	14.7
Repetitions	(D)	19.2	4.0	0	8.0

Table 2 summarises the information in Table 1. Errors due to failure to attend to all relevant aspects of the problem, incomplete correlate, are clearly the most frequent, accounting for more than half the most common errors made by people of average and above-average levels of ability. Errors due to applying the wrong principle to the solution of the problem were, if anything, most frequent at the lowest level of ability, but probably not significantly so. Errors due to confluence of ideas and repetition, not very common in any case, declined in frequency of occurrence as the total score increased.

TABLE 3

PERCENTAGE ERRORS IN POSITIONS 1 TO 8.

Position	Level of Ability			All 3 levels
	1	2	3	
1	13.5	17.6	11.4	15.1
2	10.7	13.0	12.9	12.1
3	11.9	15.5	14.7	13.9
4	15.8	14.6	10.3	14.5
5	14.5	11.5	14.7	13.1
6	7.9	7.3	7.4	7.5
7	12.4	6.1	12.1	9.5
8	13.3	14.4	16.5	14.3

Table 3 shows the distribution of erroneous choices over each position, 1 to 8, regardless of the nature of the figure chosen. If choices were made purely at random, then 12.5 per cent. of these choices should appear in each position. The fact that the pattern varies according to the level of ability is probably a function of the nature of the solutions offered, as well as their position, and appears to be related to problem-solving behaviour in general. Considering the general trend, however, it is clear that except for positions 6 and 7, the incidence of erroneous choices is about equally distributed. There is a tendency for positions 1 and 4 to be the most frequently chosen, in contrast to positions 6 and 7 which tend to be ignored. Figure 4 is close to the space to be filled on the matrix, and the popularity of figure 1 may be due to a scanning pattern. The avoidance of figures 6 and 7 appears to be related to the fact that it is people of



a fairly high level of intellectual ability who are tackling the test, since in multiple-choice tests of this nature defectives often prefer these two positions when they are reduced to random guessing. In problems 4 and 7 the figures between which choice had to be made were re-arranged to ensure a more even distribution of right and wrong choices.

By comparing the information obtained from Tables 1-3 with the curves shown in Graph 3, certain general inferences may be made about the erroneous figures chosen to complete some of the problems. In the revised Scale, problems 6, 7, 13, 17 and 24 appeared to invite mistakes at certain levels of ability. Problems 6 and 7 encouraged careless guessing long after they had ceased to be psychologically problematic. In problems 13 and 17, people appeared to be able to reduce the alternative solutions to two, well before they were able to decide which of these two was quite correct. In problem 24, one-third of the people attempting it chose the correct solution well before it reached the 60 per cent. pass level, and it was clear that in this range people were really guessing between three possible solutions. The placing of the correct solution in position 4, close to the space to be filled on the matrix may also have been the case in relatively high pass-rate on this item. This may also have been the case in problem 35 where there was a high percentage of correct guessing before the item became consistently discriminative. On the other hand, the placing of incomplete correlate figures in position 4 may have contributed to the high frequency of occurrence of this type of error on problems 26 and 32.

#### V.—CONCLUSIONS.

The results of this item analysis made it possible to eliminate twelve problems from the 1947 edition of the Advanced Matrices which contributed little or nothing to the total score obtained by adults of more than average intellectual ability. In a few test items it was possible to rearrange the figures between which choice had to be made so as to secure a more even distribution of right and wrong choices as the total score on the scale increased. In one item it was necessary to redraw the matrix of relations presented.

As far as it is possible to judge from existing data, these changes will have little or no effect on the distribution of scores obtained in forty minutes' working time. By subtracting 12 from the mean scores obtained when the former edition was given to representative groups of people, the information already published for the 1947 edition will remain relevant to that obtained with the 1962 edition.

*For the thirty-six problems retained in the new edition, the rates of change in the percentage passes on each item as the total score increases closely resemble those published by Foulds and Raven in 1950. There were surprisingly few defects in earlier editions which had not been eliminated by that time. The main defect of the 1947 edition seems to have been the rearrangement of the problems in groups of four according to the argument presented rather than their absolute order of difficulty regardless of the frequency with which they were solved. With the older edition people frequently missed problems. In the new edition the test items are placed on the right hand side of each double page only. It is strongly recommended that in future people should be told to "attempt each problem in turn and to be sure to find the correct figure to complete it before going on to the next problem which in every case becomes more difficult to solve and requires longer to check accurately." With instructions of this kind, the claim that the test assesses accurate intellectual work appears to be fully justified.*

With certain reservations the Item Analysis supports Yates' criticisms of the 1947 edition. A smaller number of items, 12 instead of 20, was found to be of poor discriminating power. This is because the range of intellectual ability included in the present parent population was intentionally chosen to include technical as well as university students. Yates confined his Item Analysis to scores obtained by university students. We did not find, as Yates did, "a very sharp increase in the percentage of subjects not attempting items after item 35," but one of his groups was only allowed 30 minutes' working time. The analysis fully confirms Yates' conclusion "that a large proportion of items 9 to 16 and 20 to 40 do discriminate successfully in relation to the total test score."

Yates also suggested that, by dropping a time limit, the test could be made one of 'power' instead of 'speed' but, in 1963, he showed that for university students limited or unlimited working time did not appear to be closely related to the prediction of examination success or failure. The 1947 edition was, in fact, designed so that it could be used without a time limit in order to assess a person's total capacity for observation and clear thinking, or with a time limit in order to assess the efficiency of his intellectual work. Until the appropriate data for the revised test are available, it is impossible to say which method of administering the test will provide the best measure of a person's intellectual performance. Omitting twelve relatively useless test items does not appreciably alter the previous recommendations. Re-arranging the problems in order of difficulty, regardless of argument, may influence its relative usefulness as a test of power as distinct from speed.

Advanced Matrices (1962) should be at least as consistent and psychologically valid as the 1947 version. The elimination of unsatisfactory items may do much to counteract any lowering of reliability due to the decrease in the number of items, but this can only be assessed as data as the new edition become available.

ACKNOWLEDGMENTS.—I should like to thank Mr. J. C. Raven for his guidance and encouragement throughout the work reported, and Dr. H. G. Bevans for his assistance during the initial stages of the enquiry.

#### VI.—REFERENCES.

- FOULDS, G. A., and RAVEN, J. C. (1950). An experimental survey with Progressive Matrices (1947), Sets I and II. *Brit. J. Educ. Psychol.*, 20, 104-110.
- RAVEN, J. C. (1941). Standardisation of Progressive Matrices, 1938. *Brit. J. Med. Psychol.*, 19, 137-150.
- RAVEN, J. C. (1948). *Progressive Matrices (1947), Sets I and II. Plan and Use of the Scale as an Individual and as a Group Test.* Obtainable from J. C. Raven.
- RAVEN, J. C. (1960). *Guide to Using the Coloured Progressive Matrices, Sets A, Ab, B.* Obtainable from J. C. Raven.
- YATES, A. J. (1961). Item Analysis of Progressive Matrices (1947). *Brit. J. Educ. Psychol.*, 31, 152-157.
- YATES, A. J. (1963). A further study of Progressive Matrices (1947). *Brit. J. Educ. Psychol.*, 33, 307-311.



# A COMPARISON OF RURAL SCHOOLS' PERFORMANCE IN A TEST-CONTROLLED AND TEST-FREE SELECTION PROCEDURE AT ELEVEN

BY H. G. ARMSTRONG  
(*Educational Psychologist, West Riding*)

**SUMMARY.** The performance at 11 of children drawn from rural schools was compared under (a) test-controlled and (b) test free conditions. There was a net increase of 7 per cent. in selective places under test free conditions and subsequent performance, as determined by progress in the grammar school, justified the increased selective provision amongst rural schools. This suggests that children in a rural environment are to some extent handicapped by an assessment which is weighted in favour of speeded tests.

## I.—DISCUSSION.

FROM general experience and previous work, there are grounds for thinking that rural children are somewhat handicapped in a selection procedure controlled by external objective tests. If so, it would follow that a larger proportion of rural children would be rightly selected by a test-free procedure. This is what has happened in the West Riding. From thirty-seven rural schools in the last three years of the test-controlled procedure, 115 children out of 863 (13·3 per cent.) were selected. From the same schools in the first three years of the test-free procedure, 192 out of 811 (23·6 per cent.) were selected. This increase of 10 per cent. represents a net gain of 7 per cent., since there was a general increase of 3 per cent. in the proportion of selective places during the interval. The subsequent performance of the rural children selected by the test-free method was on average the same as that of the urban children. The hypothesis, that the various influences to which a child in a rural environment is exposed may cumulatively produce a different tempo of life as compared with those which affect a child in an urban environment, with consequent differences in degrees of test-sophistication, is thus placed on probation.

## II.—PERFORMANCE ON (i) A TEST CONTROLLED, (ii) TEST FREE PROCEDURE.

The present investigation, covering in the first instance the four areas listed A, B, C, and D, in Tables 1 and 2 below, is designed to compare the level of selection amongst small and rural schools over a period of three years during which standardised tests controlled selection with that of the corresponding period of three years immediately after such tests were discontinued.

TABLE 1  
AGE GROUP UNDER (i) TEST-CONTROLLED PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A			842	2346
B	706	798	1268	3045
C	873	904	253	724
D	211	260	505	1463
Total	412	566		7598
	2202	2528	2868	

TABLE 2  
AGE GROUP UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	815	1026	911	2752
B	1152	1128	962	3242
C	240	236	250	726
D	488	427	438	1353
Total .....	2695	2817	2561	8073

Table 1 shows the total number of children considered for selection from all types of schools in the four areas concerned in the last three years of the tests, Table 2 the total under consideration at the end of the first three years of the new arrangements. A comparison shows that there has been an increase of 6 per cent. in the combined age group.

Tables 3 and 4 below give, respectively, the total number of children selected under the old and new schemes.

TABLE 3  
SELECTIVE PLACES GAINED UNDER (i) TEST CONTROLLED PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	139	141	148	428
B	156	242	279	677
C	60	62	61	183
D	88	108	101	297
Total .....	443	553	589	1585

TABLE 4  
SELECTIVE PLACES GAINED UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	180	178	174	532
B	269	293	282	844
C	68	76	75	219
D	145	117	125	387
Total .....	662	664	656	1982

Whilst the age group has increased by 6 per cent., a comparison between Tables 3 and 4 shows that the selected group has not risen by more than 3 per cent.



Tables 5 and 6 give, respectively, the school population amongst the small and rural schools in the areas concerned before and after the change-over.

TABLE 5

AGE GROUP IN ALL SMALL AND RURAL SCHOOLS UNDER (i) TEST CONTROLLED PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	33	26	24	83
B	59	72	74	205
C	72	61	51	184
D	57	67	74	198
Total .....	221	226	223	670

TABLE 6

AGE GROUP IN ALL SMALL AND RURAL SCHOOLS UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	42	29	31	102
B	78	73	59	210
C	55	56	66	177
D	56	58	65	179
Total .....	231	216	221	668

Table 5 shows that this group constituted 9 per cent. of the total population in the last three years of the tests whereas in the first three years of the new scheme they accounted for 8 per cent. of the population.

Tables 7 and 8 below show that 15 per cent. of the small and rural schools' population obtained grammar school places before the change and 21 per cent. after.

TABLE 7

SELECTIVE PLACES GAINED BY ALL SMALL AND RURAL SCHOOLS UNDER (i) TEST CONTROLLED PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	2	4	4	10
B	10	9	11	30
C	14	12	9	35
D	7	7	9	23
Total .....	33	32	33	98

TABLE 8

SELECTIVE PLACES GAINED BY ALL SMALL AND RURAL SCHOOLS UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	9	5	8	22
B	6	11	10	27
C	15	21	22	58
D	9	13	12	34
Total .....	39	50	52	141

The above tables do not include more than twenty-nine small and rural schools and the number has been doubled by including five more areas listed below under Table 9.

TABLE 9

AGE GROUP IN ALL SMALL AND RURAL SCHOOLS OF FIVE ADDITIONAL AREAS UNDER (i) TEST CONTROLLED PROCEDURE.

Area	No. of Small and Rural Schools	1st Year	2nd Year	3rd Year	Total
E	2	39	28	24	91
F	13	220	238	185	643
G	5	57	29	42	128
H	5	87	76	61	224
I	4	91	76	65	232
Total .....	29	494	447	377	1318

A comparison between Table 9 and Table 10 below shows that there has been a drop of approximately 20 per cent. in the small and rural school population of these areas.

TABLE 10

AGE GROUP IN ALL SMALL AND RURAL SCHOOLS OF FIVE ADDITIONAL AREAS UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
E	28	27	31	86
F	191	200	160	551
G	33	26	22	81
H	62	57	75	194
I	57	53	59	169
Total .....	371	363	347	1081



Table 11 below shows that in the areas listed 11 per cent. of the population gained selective places in the last three years of the old scheme.

TABLE 11

SELECTIVE PLACES GAINED BY ALL SMALL AND RURAL SCHOOLS IN FIVE ADDITIONAL AREAS UNDER (i) TEST CONTROLLED PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
E	5	4	0	9
F	15	19	12	46
G	14	9	6	29
H	5	14	10	29
I	9	14	7	30
Total .....	48	60	35	143

A comparison between Table 11 and Table 12 below shows that in the first three years of the new scheme, the percentage of the population obtaining selective places has risen from 11 to 16.

TABLE 12

SELECTIVE PLACES GAINED BY ALL SMALL AND RURAL SCHOOLS IN FIVE ADDITIONAL AREAS UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
E	6	5	10	21
F	16	35	22	73
G	7	6	7	20
H	13	13	11	37
I	6	5	8	19
Total .....	48	64	58	170

### III.—SCHOOLS IN A RURAL ENVIRONMENT.

The schools examined in the previous section are drawn from widely different environments. When those with urban and semi-urban backgrounds are excluded, there remains a total of thirty-seven typically rural schools which have operated a test free procedure for at least three years.

Tables 13 and 14 give, respectively, the total number of children in the age group in the last three years of the standardised tests and in the first three years of the test free procedure.

TABLE 13

RURAL SCHOOLS' AGE GROUP IN ALL AREAS UNDER (i) TEST CONTROLLED PROCEDURE.

Area	No. of Rural Schools	1st Year	2nd Year	3rd Year	Total
A	5	33	26	24	83
B	3	11	9	12	32
C	13	72	71	51	194
D	5	38	42	57	137
F	7	94	114	78	286
G	1	6	7	5	18
H	2	41	33	24	98
I	1	2	9	4	15
Total .....	37	297	311	255	863

TABLE 14

RURAL SCHOOLS' AGE GROUP IN ALL AREAS UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	42	29	31	102
B	9	16	14	39
C	55	56	66	177
D	42	35	47	124
F	90	89	79	258
G	7	3	0	10
H	29	27	37	93
I	4	2	2	8
Total .....	278	257	276	811

Tables 15 and 16 below gives, respectively, the total number of selective places obtained in the last three years of the old and the first three years of the new scheme.

TABLE 15

SELECTIVE PLACES GAINED BY RURAL SCHOOLS IN ALL AREAS UNDER (i) TEST CONTROLLED PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	2	4	4	10
B	4	2	4	10
C	14	12	8	34
D	3	4	8	15
F	8	12	4	24
G	3	2	0	5
H	3	7	6	16
I	0	1	0	1
Total .....	37	44	34	115



TABLE 16

SELECTIVE PLACES GAINED BY RURAL SCHOOLS IN ALL AREAS UNDER (ii) TEST FREE PROCEDURE.

Area	1st Year	2nd Year	3rd Year	Total
A	9	5	8	22
B	1	2	6	9
C	15	21	22	58
D	8	8	10	26
F	10	22	20	52
G	2	0	0	2
H	5	6	7	18
I	2	2	1	5
Total .....	52	66	74	192

Out of a total of 863 children in the one case, 115 gained selective places and from a total of 811 in the other, 192 obtained grammar school places. In a test-controlled selection procedure 13.3 per cent. and in a test-free selection procedure, 23.6 of the age group from rural schools gained grammar school places.

IV.—SUBSEQUENT PERFORMANCE UNDER A TEST-FREE PROCEDURE OF (a) ALL SMALL AND RURAL SCHOOLS, (b) RURAL SCHOOLS ONLY.

The tables in Sections II and III show that the number of children admitted to grammar schools from small and rural schools has increased and the details listed under Tables 17 to 19 below throw some light on the later progress of the increased numbers selected.

A rank order is obtained from the grammar school at the end of the first year course for the complete entry, in order to determine what proportion of an individual school's entry falls in the upper two-thirds and lower third of this order. This information is an integral part of the test-free procedure now in operation and constitutes a regular 'feed-back' each year from the secondary to the primary schools.

TABLE 17  
FOLLOW-UP OF ALL SMALL AND RURAL SCHOOLS.

Area	1st Year		2nd Year		3rd Year		Total	
	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3
A	7	—	2	1	4	3	13	4
B	2	2	6	2	3	3	11	7
C	10	5	12	8	11	7	33	20
D	1	3	4	3	4	3	9	9
E	5	1	1	3	—	—	6	4
F	4	5	22	7	—	—	26	12
G	3	2	2	2	—	—	5	4
H	8	2	2	3	—	—	16	5
I	3	1	3	2	—	—	6	3
Total	43	21	60	31	22	16	125	68

Table 17 gives, respectively, three consecutive entries from four areas and two from five areas. In area A for instance, all seven children admitted are to be found in the upper two-thirds of the complete entry at the end of the first year in the grammar school. Of the total number admitted over a period of three years, thirteen are placed in the upper two-thirds and four in the lower third of the total entry. From all areas, out of a total of 193 children admitted, 125 are to be found in the upper two-thirds and sixty-eight in the lower third of the entry.

TABLE 18  
FOLLOW-UP OF RURAL SCHOOLS ONLY.

Area	1st Year		2nd Year		3rd Year		Total	
	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3
A	7	—	2	1	4	3	13	4
B	1	—	1	1	2	2	4	3
C	10	5	12	8	11	7	33	20
D	1	3	3	3	3	3	7	9
F	1	3	15	5	—	—	16	8
G	2	—	—	—	—	—	2	—
H	2	1	4	1	—	—	6	2
I	—	1	1	1	—	—	1	2
Total	24	13	38	20	20	15	82	48

Table 18 above gives the same information as Table 17, with one area eliminated because of its urban and semi-urban environment. Of the total number of children involved, namely 130, eighty-two are placed in the upper two-thirds of the order and forty-eight in the lower third. Both Tables show clearly that the children have justified their selection.

TABLE 19  
LONG TERM FOLLOW-UP.

Area	No. of Schools	1st Year		2nd Year		3rd Year	
		Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3	Upper 2/3	Lower 1/3
A	5	7	—	4	—	—	—
B	3	3	2	3	2	5	—
C	13	10	4	9	4	3	—
D	5	1	3	2	2	11	2
Total	26	21	9	18	9	22	3

Table 19 above shows the position of the entries from twenty-six rural schools after three years in the grammar school. In area A, for example, of the seven children admitted and placed in the upper two-thirds of the order at the end of the first year, five are found to be in the same position at the end of



three years in the grammar school. Similarly in area C, of the fourteen children admitted and divided as indicated after one year, eleven are to be found in the upper two-thirds of the order and two in the lower third three years later. The data show that the children from these twenty-six rural schools not merely maintain but improve their position as they move up the school.

#### V.—NOTE ON A TEST-FREE PROCEDURE.

This procedure, which has superseded standardised tests, is based on the assumption that past performance is a reasonably good guide to future success, and that, whilst the proportions sent forward to grammar schools may vary considerably between one school and another, they do not alter greatly for any one school from one year to the next. In the initial research which was undertaken before the scheme was launched, the number of grammar school places obtained by a typical school was found to be within four of its average, that is to say, the standard deviation of places gained by a school seldom exceeded four.

The scheme was introduced experimentally for two years in a relatively simple part of the County, after which time it replaced the operative scheme, and has continued to spread progressively throughout the area. At the present time, it covers two-thirds of the administrative area.

When the scheme is introduced for the first time into a new area, the primary schools are asked to rank their children in order of merit, using as a broad criterion suitability for a grammar school course. Each school is given a provisional quota based on the average number of places obtained during the past three years. The last two children on the schools' provisional list, together with the next two below them, constitute a borderline group for the individual school. These four children and other groups of four from an appropriate maximum of eight schools, are seen by a panel of five experienced teachers, three of whom are drawn from primary schools and two from secondary schools.

The function of the panel is to fill the sixteen places from the thirty-two children they meet on the first occasion. If, however, all four from one school take up places, the next four below them are seen at a later date. The same is true for a school that obtains none, but in this case it is the next four above who are seen. This continues until a split is made in a school's list, and the final border-line drawn for one school as compared with another. About 25 per cent. of the schools in an area require a second round and fewer than 1 per cent. a third round.

The children spend a school day with the panel and, although no strict rules determine what is done, generally the rhythm of a primary school day is followed with written assessments in the morning and oral and individual work in the afternoon. That variations in procedure exist between different areas is fully acknowledged but, since the provisional quotas ensure that the children who appear in the first instance are near the border-line, uncontrolled variation is eliminated.

Each year information is fed back to the primary schools on the progress of the children admitted to a grammar school. This provides a check on the efficiency with which the panels maintain the correct borderline between one school and another. By dividing the complete entry at the secondary stage into an upper two-thirds and a lower third, it is possible to enumerate the primary schools whose border-lines have been drawn either too high or too low. In a recent follow-up of over two hundred primary schools, less than 5 per cent. appeared to have had their border-lines wrongly drawn. If later progress is a valid criterion, then the efficiency of the scheme is unquestioned.

In each area where it operates, there is a varying proportion of schools with a small number of children in the age group. These may, like the bigger schools, send to the panel up to four children on the first occasion, but each is required to send at least one in order to control any variation in judgment that may arise between one small school and another.

#### VI.—GENERAL OBSERVATIONS.

Researches, similar to the one reported here, have hitherto been carried out at second-hand. In a recent investigation by Moreton and Butcher, the effect or influence of speeded tests on rural children was examined with reference to performances at 11. The investigation was started because one of the authors met with frequent complaints from teachers and parents that the country child is at a disadvantage in the working of speeded tests at 11+. This disadvantage is often attributed to a slower tempo of life in the country as compared with that in urban areas.

Children from a working class area of central Manchester were compared with children from a small rural area of Westmorland, on moderately speeded tests of Arithmetic, English and Verbal Reasoning, as used in an 11+ procedure. Manchester scores were slightly higher in all three papers, two of the three differences being statistically significant. The hypothesis, that speeded tests to some degree handicap rural children was considered to be confirmed, but the relative differences were estimated to be so small as to suggest that no change be made in the method of selection.

From the present research, it is clear that where standardised tests play no part in the assessment of children from rural schools, they not only have a greater opportunity of obtaining grammar school places, but also merit their selection if subsequent performance is an acceptable criterion.

With the changes at present taking place in the organisation of secondary education, it should be possible to subject the hypothesis to still further tests.

#### VII.—REFERENCES.

- MORETON, C. A., and BUTCHER, H. J. (1963). Are rural children handicapped by the use of speeded tests in selection procedures? *Brit. J. Educ. Psychol.*, XXXIII, 1.

(Manuscript received 15th November, 1963)



# DATE OF BIRTH, BACKWARDNESS AND EDUCATIONAL ORGANISATION

By PHILLIP WILLIAMS

(University College of Swansea Department of Education)

**SUMMARY.** An examination of previous work on the relation of season of birth to educational backwardness suggests that an undue preponderance of Summer-born children might be found in special schools for the educationally subnormal. This suggestion is verified for a population of 265 children drawn from such schools. Two possible reasons for this effect are considered, that of differential entry to infant school, and the 'age-group-position' effect. It was not found possible to distinguish these effects in this study. The results of this study are collated with those of previous investigations into the relationship of season of birth and various forms of handicap to show the relative consistency of different evidence. Increased flexibility of educational organisation and the extension of services for the educational guidance of individual children are advocated.

## I.—INTRODUCTION.

THERE is accumulating a body of evidence on the relation of season of birth to educational progress. Thirty years ago, Pintner and Forlano<sup>7</sup>,<sup>8</sup> found a tendency for higher scores on standardised tests of ability to be linked with Summer season birthdays. They showed this to be independent of the calendar months, which fall in different seasons, according to geographical hemisphere, and assumed that the effect was due to climatic influences in the first few months of growth.

Fitt<sup>2</sup> considered the question of the influence of season of birth from a wider standpoint. After a careful review of the evidence, he concluded that the brighter, the stronger, the heavier, etc., tended to be born in the Summer/Autumn seasons. He ascribed this effect not to seasonal influences at or immediately succeeding birth, but to seasonal influences on conception. Roberts<sup>10</sup> while not disputing Fitt's evidence, puts forward an alternative explanation of the findings in relation to intelligence.

More recently, Orme<sup>4</sup> has demonstrated that in a group of severely subnormal individuals, lower intelligence is more frequently linked with Winter and Spring births. He has also suggested (Orme<sup>5</sup>) that climatic temperature changes occurring in pregnancy may be one factor underlying the relationships found between season of birth and subsequent intellectual performance.

The interest of this work lies in its relation to findings in the field of education, which indicate the disadvantage of Summer birth. John<sup>3</sup>, for example, in an article based on a paper read to the 1959 meeting of the British Association for the Advancement of Science, showed that among 1,300 retarded readers there was a significantly higher proportion of children with Summer birthdays than Autumn ones, whereas the proportion of children with Spring birthdays approximated to that expected.

Williams<sup>13</sup> confirmed this finding for a group of children who were very retarded in reading, and for a group of 342 educationally backward children. He also showed that in a group of 1,511 children referred for child guidance, Summer birthdays again predominated significantly. Pidgeon and Dodds<sup>6</sup> working not with a selected group of children who were handicapped educationally or emotionally, but with a full range of children attending six junior schools

found a positive correlation between scores on a reading test and length of schooling, when the effect of age was eliminated. The relationship was significant for older children in two schools which had been streamed by age.

These last three investigations show the educational and emotional disadvantages—one might almost say the developmental disadvantages—of Summer birth. Explanations so far suggested link mainly with the question of length of stay in infant schools which can be called the 'term-of-entry' effect. In the areas in which the above investigations were carried out, infant schools accepted three intakes of children each year—in September, January and April. The children transferred to junior schools once a year, an arrangement which results in varying lengths of infant education for children depending on their birthday. This is shown in Table 1 :

TABLE 1  
EFFECT OF BIRTHDAY MONTH ON INFANT SCHOOL STAY.

Birthday Month ...	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
Enters Infants ....	Jan.	Jan.	Jan.	Apr.	Apr.	Apr.	Apr.	Apr.	Sept.	Sept.	Sept.	Sept.
No. of terms in Infants .....	8	8	8	8	7	7	7	7	9	9	9	9

There is a side effect here which penalises the Summer-born child again. Not only does he spend less time in the infant school, but he joins the school at a time when the school numbers are at their annual peak, and his class is likely to be larger than that which the Autumn-born child entered two terms earlier.

There is, however, a second possible explanation of the findings. Children who are born in the Summer months usually retain their position of youngest in their age group throughout the junior school year, where promotion tends to be by age. This poses additional stress since teaching levels will tend to be harder for younger children. Thus, where it is found that Summer-born children are at a disadvantage in comparison with Autumn-born children, the effect may be due not to the length of schooling, but to the related factor of the position in the age group. Or both factors may be operating. (Pidgeon and Dodds (*op. cit.*)) suggest a third factor which may operate in certain circumstances. They suggest that where streaming by age is carried out, the effect may be accentuated, since the older stream gives the impression of being the 'better' stream and is, therefore, 'advanced'.)

These factors link with the nature of the educational system, and with the administration of education. The disadvantages of Summer birth in this country are not necessarily opposed to the earlier evidence of the advantages of Summer birth, much of which is based on work in countries with different educational arrangements, and which relates to other variables than attainment.

## II.—THE PRESENT INVESTIGATION.

The Department of Education at Swansea is currently engaged in a research project in special education, supported financially by the Ministry of Education. Seven local education authorities in South Wales are co-operating in the project, and have made available records of children attending the eight special schools for the E.S.N. which they maintain. Six schools are day, and two residential, and data relating to 265 children have been obtained in this section of the research project.



Since in all the seven authorities promotion from a lower section of the school system to a higher one is usually based on age on 31st August or 1st September, it seemed reasonable to hypothesise that there would be a higher proportion of Summer-born children than Autumn-born children in this population, for the following reasons. One of the factors in selection for special schools for the E.S.N. is low educational performance allied to low scores on individual assessments of intellectual development. The Chief Medical Officer of the Ministry of Education<sup>1</sup> reports that over the ten-year period 1950-1960, the waiting list for places in special schools for the E.S.N. remained fairly constant at 12,000, although the total number of school places provided doubled. Where places are limited, as is nearly always the case, lower educational performance results in a greater likelihood of selection. Lower educational performance is also related to Summer birthdays.

The 265 children are made up of three groups: (i) a one-in-five sample by register order of all 800 children on the books of the eight schools at the start of the Autumn term, 1962 (N group); (ii) a group of fifty-seven children who, during the course of a five-year period from 1st September, 1957, to 31st August, 1962, were transferred from special schools to the ordinary school system, usually secondary schools (S group); (iii) a group of forty-seven children who, during the same five-year-period were transferred from the special schools to occupation centres (now training centres) (E group).

Data relating to all three groups of children are now being analysed, and a clinical investigation is being made of the last two groups.

There might also be a case for a marginal effect on intelligence test scores themselves. Vernon<sup>12</sup> has shown the effect of educational environment on group test scores, and it would seem reasonable to expect a similar sort of effect with individual tests of the type used in selection for special education. So it would seem likely that the Summer birthday effect might appear with the E.S.N. children from this project.

### III.—RESULTS.

The distribution of birthdays of these 265 children is given below in Table 2:

TABLE 2  
BIRTHDAY DISTRIBUTION, BY MONTH.

Month .....	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
No. ....	26	19	21	22	26	21	27	37	10	19	17	20

Previous work (John, Pidgeon and Dodds, *op. cit.*) suggests that the main cause of the Summer birthday effect may be the differential effect of termly entrance to infants' schools. If this is so, then the frequencies obtained by condensing Table 2 into the appropriate three categories should be significantly different from those expected. This is done in Table 3, which also shows the expected frequencies.

The frequencies expected can be calculated from the figures available for seasonal birthrates (Registrar-General)<sup>9</sup>. Although there is not a very great variation between the birthrate at different times of the year, the birthrate is at a minimum in November, rises to March, remains high to May or June, and then declines again, except for a minor rise in September. But there have been some variations in this picture, and in order to obtain fair estimates for the expected frequencies, the data given for the years 1944-1955, the years in which nearly all the children in the group were born, were used. Estimated in this way, the frequencies of birthdays in the three categories are given in Table 3.

TABLE 3  
BIRTHDAY DISTRIBUTION, BY TERM.

Term of Birth	Autumn (Sept.—Dec.)	Spring (Jan.—April)	Summer (May—Aug.)	Total
N (observed) . . . .	66	88	111	265
N (expected) . . . .	83	91	91	265

$$\chi^2 = 7.98. \text{ For 2 d.f., } 0.01 < p < 0.02.$$

This distribution of births is clearly significantly different from that obtained in the population as a whole.

If, in view of this it is accepted that chance factors are unlikely to be responsible for this distribution of births, it is equally not necessary to assume that educational organisation is responsible. It is not impossible that there are congenital factors operating which make it more likely for Summer-born children to be slow learners than Autumn-born children. Stott<sup>11</sup> has pleaded the case of pre-natal influences as causative factors in the production of backwardness, and it is possible that the season of pregnancy may affect some of these influences. But it seems reasonable to consider the distribution to be linked with educational organisation, since: (a) there does not appear to be a logical alternative hypothesis at present; (b) the distribution agrees with the realities of the educational year—it has a minimum in September and a maximum in August, which are the months of starting and ending the educational year in all seven local education authorities; (c) theories linked with educational organisation have been suggested for similar effects with other groups of children.

If we accept that this distribution is likely to be linked with educational organisation, it remains to be considered which of the two main possible explanations, the 'term of entry' effect, or the 'age-group position' effect is operating.\* One way of distinguishing these two factors is by considering the

\* Streaming by age is rare in the area of the investigation and is unlikely to have affected more than a handful of children at the most.



distribution of frequencies within the terms. If a 'term of entry' effect is operating then the frequency distribution could be expected to approximate to a series of steps, with plateaux over the terms and rises between them. If an age-group position effect is operating then the frequency distribution is probably linear.

Or both effects could be operating, producing an intermediate picture.

If the months of the year are reclassified by their positions within the terms, taking September, January and May as first months, October, February and June as second months, etc., Table 4 can be drawn up.

TABLE 4  
BIRTHDAY DISTRIBUTION, BY MONTH POSITIONS.

	First month	Second month	Third month	Fourth month
Autumn Term ...	10	19	17	20
Spring Term ....	26	19	21	22
Summer Term ...	26	21	27	37
N .....	62	59	65	79

The term-of-entry effect is now held constant, and if the age-group-position effect is operating, the frequencies obtained for the first months should be below that expected, the frequencies for the fourth months should be above that expected, and intermediate frequencies should fall in appropriate positions. The expected frequencies can again be calculated from the Registrar-General's figures, and Table 5 obtained.

TABLE 5  
OBSERVED AND EXPECTED BIRTHDAY DISTRIBUTION, BY MONTH POSITIONS.

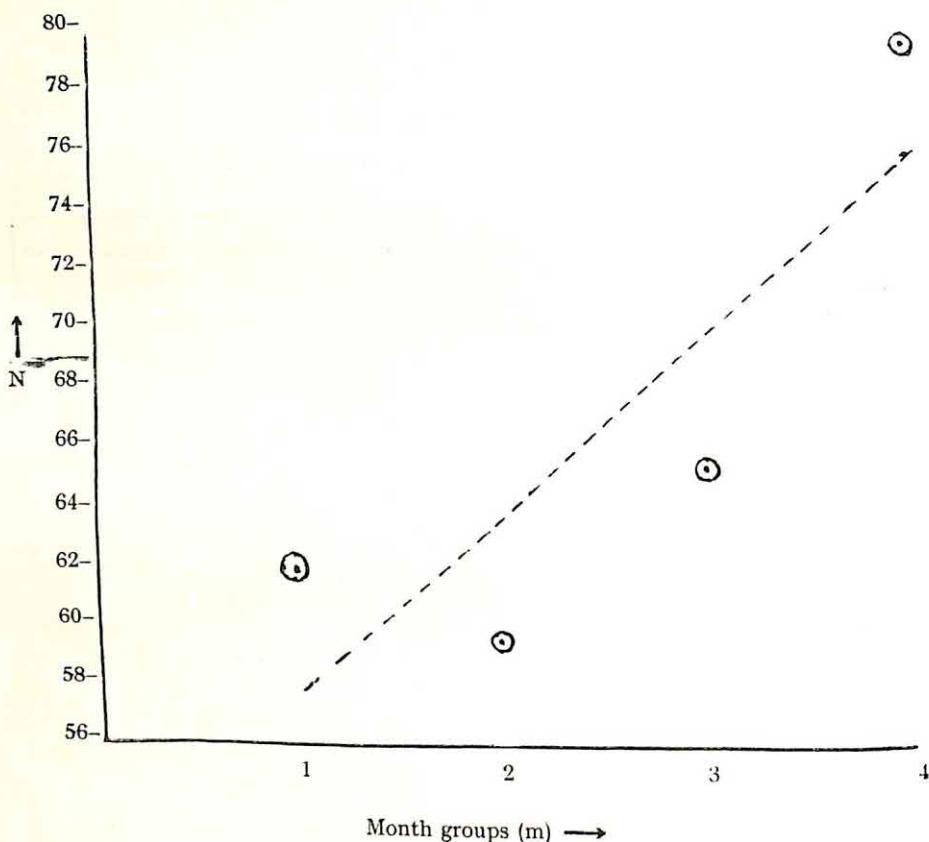
	First month	Second month	Third month	Fourth month
Expected .....	68	65	67	65
Observed .....	62	59	65	79

The picture corresponds with that expected if an age-group position effect is operating, but  $\chi^2 = 4.16$ . For 3 d.f.  $0.20 < p < 0.30$ .

The presence of a term-of-entry effect remains to be considered. An estimate can be made by attempting to allow for the age-group position effect.

If we assume that this is linear, then the best straight line that can be fitted to the points in diagram 1 is  $N = 52 + 5.7m$ .

DIAGRAM 1

BIRTHDAY DISTRIBUTION *v.* MONTH GROUP

The slope, 5.7, represents the average frequency increase between the month groups, which differ from each other by the effect of three months. The best estimate of the age-group position effect would, therefore, seem to be a frequency increase of 1.90 per month.

This estimate of the age-group position effect can now be used to weight the observed frequencies, in order to see if there is a term-of-entry effect remaining.



The year can be split at the February/March division, and the month frequencies can be weighted in accordance with their distance from this zero point.

TABLE 6  
WEIGHTED BIRTHDAY DISTRIBUTION, BY MONTH.

Month .....	September	October	November	Dec.	January	February
Observed .....	10	19	17	20	26	19
Weight .....	$\frac{+11 \times 1.9}{2}$	$\frac{+9 \times 1.9}{2}$	$\frac{+7 \times 1.9}{2}$	$\frac{+5 \times 1.9}{2}$	$\frac{+3 \times 1.9}{2}$	$\frac{+1.9}{2}$
Weighted frequency	20.45	27.55	23.65	24.75	28.85	19.95
	March	April	May	June	July	August
Observed .....	21	22	26	21	27	37
Weight .....	$\frac{-1.9}{2}$	$\frac{-3 \times 1.9}{2}$	$\frac{-5 \times 1.9}{2}$	$\frac{-7 \times 1.9}{2}$	$\frac{-9 \times 1.9}{2}$	$\frac{-11 \times 1.9}{2}$
Weighted frequency	20.05	19.15	21.25	14.35	18.45	26.55

The weighted frequencies can now be compared with the expected frequencies by terms, in order to see if a term-of-entry effect remains.

TABLE 7  
BIRTHDAY DISTRIBUTION, WEIGHTED AND EXPECTED, BY TERMS.

Term	Autumn (Sept.—Dec.)	Spring (Jan.—Mar.)	Summer (April—Aug.)	N
Weighted frequencies	96.4	88	80.6	265
Expected frequencies	83	91	91	265

This distribution runs counter to that expected :

$$\chi^2 = 3.45. \text{ For 2 d.f., } .10 < p < .20.$$

#### IV.—DISCUSSION.

It seems fairly clear that for the 265 children who have been attending special schools for the E.S.N. in South Wales, a disproportionate number have Summer birthdays. It would also seem reasonable to suppose that this is an effect of educational organisation, although other factors may be operating.

While this problem may not bulk very large in schools dealing with a full range of children, it has a considerable impact in services dealing with handicapped children of various kinds. Some of the evidence from previous investigations is collated with this in Table 8.

TABLE 8

SEASONAL BIRTHDAY DISTRIBUTION OF HANDICAPPED CHILDREN—FIGURES AS PERCENTAGES.

Group Investigated	Ref.	Season of Birth			N
		Autumn	Spring	Summer	
Retarded Readers . . . .	3	27	32	41	1600
Retarded Readers . . . . .	13	27	35	38	382
Child Guidance Cases . .	13	28	35	37	1511
E.S.N. children . . . . .	This study	25	33	42	265
" B " and " C " Classes	14	28	34	38	729

The appearance of this effect in child guidance cases is intriguing, but not necessarily inconsistent with the general theme of increased educational difficulties of Summer-born children. Emotional and educational difficulties in children are not the same, but they are often linked. Where one group of children is educationally handicapped, as appears to be the case with Summer-born children, it is reasonable to expect that the increased stresses will result in a greater proportion of emotional problems in that group, and a greater frequency of referrals to child guidance clinics. Perhaps more important than these side-effects are the possibilities of other side effects in associated fields, possibilities to which these findings give rise. One wonders also if Autumn-born children predominate as markedly at the other end of the educational spectrum.

It is clearly important to establish the cause or causes for this effect. Where a Summer-born child stands nearly a 50 per cent. greater chance of needing some form of special help than an Autumn-born child, there is need for modification of some sort. In this investigation it has not been possible clearly to separate the age-group-position effect from the term-of-entry effect. The linear nature of the age-group-position effect rests on an assumption which, although reasonable, may not be borne out by further enquiry. If a definite choice had to be made, then for these children, the age-group position effect would seem to be more likely to be linked with the observed birthday distribution than the term-of-entry effect. But further investigations would be necessary to support or contradict this.

It is also possible that different effects are responsible for the same type of distribution in different groups. Thus, with children who have been attending special schools for the E.S.N., it is possible that the start of schooling was delayed, so as to affect the issue as examined here. With other groups of children, e.g., child guidance cases—this is unlikely to be the case. The question is clearly a complex one, but its implications for the needs of individual children are quite definite. More flexible forms of educational organisation, more attention within the ordinary schools to the extra problems of Summer-born



children, and especially, perhaps, the extension of services providing guidance for individual children early in their school careers, are all remedies which need to be tried.

ACKNOWLEDGMENTS.—The writer wishes to thank the Ministry of Education for its financial support of the investigations of which these findings are a part. He is grateful to the Directors of Education and the Medical Officers of Health of the Counties of Cardiganshire, Carmarthenshire, Monmouthshire, Pembrokeshire, the County Boroughs of Newport and Swansea, and the City of Cardiff for giving permission for these investigations to proceed. He wishes to record his appreciation of the co-operation given by the Heads of the various schools, Mr. D. Burn, Mr. E. V. Earl, Mr. T. D. Hopkins, Miss N. Jones, Mr. R. N. Lewis, Mr. H. Llewellyn, Mr. E. Scriven, and Mr. J. Sloman. He is grateful to Mr. I. G. Evans and Dr. S. K. Zaremba for their help with the data analysis, to Mr. R. R. Dale for his useful comments, and to Miss M. Lewis, the project secretary.

#### V.—REFERENCES.

- <sup>1</sup> CHIEF MEDICAL OFFICER OF THE MINISTRY OF EDUCATION (1962). *Health of the School Child, 1960-1961*. H.M.S.O.
- <sup>2</sup> FITT, A. B. (1941). *Seasonal Influences on Growth, Function and Inheritance*. Oxford: Oxford University Press.
- <sup>3</sup> JOHN, E. (1962). *The Age Factor in Reading Retardation—Researches and Studies*. University of Leeds Institute of Education, No. 24, pp. 1-7.
- <sup>4</sup> ORME, J. E. (1962). Intelligence and season of birth. *Brit. J. Med. Psychol.*, 35, 233-234.
- <sup>5</sup> ORME, J. E. (1963). Intelligence, season of birth and climatic temperature. *Brit. J. Psychol.*, 54, 3, pp. 273-276.
- <sup>6</sup> PIDGEON, D. A., and DODDS, E. M. (1961). Length of schooling and its effect on performance in the junior school. *Educ. Research*, III, pp. 214-221.
- <sup>7</sup> PINTNER, R. and FORLANO, G. (1933). The influence of month of birth on intelligence quotients. *J. Educ. Psychol.*, 24, 561-584.
- <sup>8</sup> PINTNER, R., and FORLANO, G. (1939). Season of birth and intelligence. *J. Genet. Psychol.*, 54, 353-358.
- <sup>9</sup> REGISTRAR-GENERAL (1962). *Statistical Review of England and Wales for 1960, Pt. II*. H.M.S.O., and for years 1944-1955.
- <sup>10</sup> ROBERTS, J. A. F. (1944). Intelligence and season of conception. *Brit. Medical J.*, 4th March, 1944, pp. 320f.
- <sup>11</sup> STOTT D. H. (1963). How a disturbed pregnancy can harm the child. *New Scientist*, 320, pp. 13-17.
- <sup>12</sup> VERNON, P. E. (1960). *Intelligence and Attainment Tests*. University of London Press, p. 154.
- <sup>13</sup> WILLIAMS, P. (1959). *The Growth of Reading Vocabulary*. Unpublished Ph.D. Thesis, University of London Library.
- <sup>14</sup> JINKS, P. C. (1964). An investigation into the effect of date of birth on subsequent school performance. *Educ. Research*, VI, No. 3, pp. 220-225.

(Manuscript received, 20th December, 1963)

# THE SOCIAL AND EDUCATIONAL BACKGROUND AND ANTICIPATED CAREER PROSPECTS OF A GROUP OF STUDENTS IN A COLLEGE OF ADVANCED TECHNOLOGY

By RUTH RICE

SUMMARY. The Robbins' Committee on higher education concluded: "We envisage a new role for the Colleges of Advanced Technology as technological universities, conferring their own degrees."<sup>1</sup> They reassert that "The Diploma in Technology is the equivalent of an honours degree."<sup>2</sup> The purpose of the survey, conducted in 1961, concurrent with many of the enquiries for the Robbins' Report, on which the following extract is based, was to throw light on the social and educational background and career prospects of students studying for a Diploma of Technology in a College of Advanced Technology. Certain conclusions are drawn, viz., that this College attracts a predominately local population of students, the majority of whom, under our present educational system, would not find their way to the university but who, upon qualification, will have the same chance as the university graduate to reach a position of status in industry. As many changes are taking place in this field of higher education, it is hoped that the findings of this survey will be of interest to the social historian in comparing the students who are pioneering this qualification in industry with those who succeed them.

## I.—INTRODUCTION.

As the First and Second Progress Reports of the Nuffield Research Project, conducted by Dr. Marie Jahoda of Brunel College of Technology were available, it was possible to follow up many of the themes of her work. Similar questions were asked on the student's academic and social background; how he compares the Diploma of Technology with a B.Sc.Tech.; his doubts about his career; the reasons for promotion; the amount he hopes to earn at the end of his course, five years later, and at the height of his career. Dr. Marie Jahoda's work at Brunel College of Technology was a vanguard study to these investigations.

As the number of students was small, it was decided to include them all rather than to draw a sample. The questionnaire was given to all 1st, 2nd, 3rd and 4th year students of the Mechanical and Electrical Engineering Departments and to the 1st, 2nd and 3rd year of the Building and Civil, and Chemical Engineering Departments. The former two departments have a twice-yearly intake of students in October and January; the latter two departments take students in October only and had at the time of the survey no 4th year students.

From a pilot survey conducted previously it was evident that all students had much in common; the disparity of numbers between departments made comparison impossible except in the case of the Electrical and Mechanical Engineering Departments.

Wherever possible, students were asked to fill in the questionnaire during a one-hour lecture period. An attempt was made to make the atmosphere and questionnaire as informal as possible, although a member of staff was present to prevent collaboration between students. Where it was not possible to make use of a lecture period a questionnaire was posted to the student's home. It was made clear that the completion of the questionnaire was entirely voluntary. By dispensing with personal interviews and assuring the students that their replies would be treated anonymously, it was hoped that students would be encouraged to make forthright and honest replies. Of the 465 Diploma of



Technology students enrolled in the four departments during the academic year 1960-61, 392 (84 per cent.) completed the questionnaire. The main reason given for not completing the questionnaire was that the student was too busy with other college work. In this report it is not possible to give the detailed information contained in the tables of the survey. In selecting significant material, I have included comments based on my own experience of working and teaching in the College for seven terms.

TABLE 1

AN ANALYSIS OF STUDENT INFORMANTS BY DEPARTMENT AND YEAR OF COURSE.

QUESTION : Give date (month and year) of intake.

Please state the year of your course and the name of your department.

All Student Informants by Department and Year of Course		Number
Electrical Engineering Department .....	1st year .....	58
	2nd year .....	49
	3rd year .....	38
	4th year .....	28
Mechanical Engineering Department .....	1st year .....	69
	2nd year .....	51
	3rd year .....	31
	4th year .....	22
Building and Civil Engineering Department .....	1st year .....	16
	2nd year .....	7
	3rd year .....	4
Chemical Engineering Department .....	1st year .....	4
	2nd year .....	6
	3rd year .....	9
All Students .....	Total .....	392

## II.—FROM WHERE THE COLLEGE DRAWS ITS STUDENTS.

The Robbins' Committee drew a distinction between Regional Colleges and Colleges of Advanced Technology. "Taking the two groups of colleges as a whole, the most significant difference from our point of view is that Colleges of Advanced Technology are already national institutions—firms of national repute tend to gather together in one centre the students whom they wish to send on sandwich courses."<sup>3</sup> Referring to Regional Colleges it states, "with a few exceptions their area of recruitment is largely regional."<sup>3</sup> In contrast to this general distinction, the majority of Diploma of Technology students at this College of Advanced Technology are drawn from the surrounding conurbation, although a slightly higher proportion of students (62 per cent. to 54 per cent.) live with their parents during the work period than do during the College period. There is no evidence at present that many students from further afield wish to attend the College. The College is situated centrally in regard to several large industrial organisations which account for most of the sponsored students. Firms—even those on a national scale—are not likely to send sponsored students to Colleges outside their immediate vicinity and, therefore, the College of Advanced Technology will attract a predominantly regional population.

## III.—THE STUDENT'S SOCIAL BACKGROUND.

TABLE 2

THE SOCIAL CLASSIFICATION OF THE FATHER'S OCCUPATION.

QUESTION : What does your father do for a living (give detailed specification) ?

Status of Father's Occupation	Total
Base of %—All Students . . . . .	392
Proportion who were classified as :	%
AB . . . . .	22
C . . . . .	29
D <sub>1</sub> . . . . .	25
D <sub>2</sub> E . . . . .	17
Unclassified/no father . . . . .	7

If the father's occupation is taken as a gauge of the student's social background, it will be seen from Table 2 that this College of Advanced Technology draws students from the whole gamut of our class structure although only 17 per cent. of children of lower working class (D<sub>2</sub>E) parents attend as opposed to 29 per cent. from middle-class homes (C). No claim is made for the complete accuracy of this table as some students did not give sufficiently detailed specification of what their fathers do—in particular the word 'engineer' without qualification was almost impossible to classify. This was the occupation of several fathers and shows the direct, if sometimes unrecognised, influence of the father on the student's choice of career. The selection of students based on ability alone rather than considerations of class, plus the fact that 64 per cent. of the students are able to support themselves fully while studying, indicates an advance towards the ideal of equality of opportunity in further education.

## IV.—THE MARITAL STATUS OF STUDENTS.

Of the 392 students interviewed, 5 per cent.—all men—were married. There seemed much disagreement as to whether marriage would favourably affect the student's attitude to work. Several students were unwilling to express an opinion. 18 per cent. thought that marriage would give an incentive to qualify, while 13 per cent. thought that students would be helped by being settled with their own homes. Less time for study and more worries were the main reasons why students thought marriage would unfavourably affect their attitude to work. 12 per cent. expressly stated that marriage would reduce the time for study. Most students feel free to marry and, therefore, no burning interest was shown in this question.

## V.—THE STUDENT'S EDUCATIONAL BACKGROUND.

*The type of school last attended full-time.* The educational background of the students varied considerably. 62 per cent. of the 392 students interviewed had attended a grammar school full-time, 16 per cent. a technical school, 11 per cent. a public school and 8 per cent. a secondary modern school. The Mechanical Engineering Department (173 students interviewed) had a higher proportion of secondary modern school students (12 per cent.) than the Electrical Engineering



Department (173 students interviewed) (5 per cent.) and a lower proportion of grammar school students (59 per cent. : 65 per cent.). It seems that the proportion of technical school students is increasing (20 per cent. of the 147 1st year students as against 12 per cent. of the 113 2nd year and eighty-two 3rd year) unless this difference can be accounted for by the yearly wastage rate.

*Age on leaving full-time day school.* The majority of students did not leave school before 1955. There were fewer mature students on the Diploma of Technology course than might be expected on the grounds that previous industrial experience is desirable and 64 per cent. of students said they were able to fully support themselves. Nevertheless there were considerable differences in the age at which students had left full-time day school. 15 per cent. left under the age of 16 years, while 38 per cent. left before their seventeenth birthday. It is unlikely that under our present educational system these students would have found their way to the university. 8 per cent. left school at 19 years and over. From replies given to other questions it seems reasonable to assume that some of these students had failed to gain university entrance.

A higher proportion of students of the Mechanical Engineering Department had left school at a younger age than students of the Electrical Engineering Department—46 per cent. compared with 32 per cent. left school under the age of 17 years (20 per cent. compared with 11 per cent. left under 16 years). Likewise, a higher proportion of Electrical Engineering students had remained at school over the age of 18 years—51 per cent. compared with 36 per cent. of Mechanical Engineering students. The ages of the students on leaving school do not vary greatly over the four years interviewed.

*Attendance at Day Release and Evening Classes.* Many students had supplemented their full-time school education before coming to this College of Advanced Technology. 56 per cent. had been day release students and 44 per cent. had attended evening classes.

Of the 221 students who had been on day release schemes, 22 per cent. attended for under 1 year ; 39 per cent. for 1-2 years ; 34 per cent. 3-4 years, and 5 per cent. 5 years and over. In order to find out the general opinion of students on the advantages and disadvantages of day-release training as compared with full-time study, all students were asked to comment. 57 per cent. declined to suggest any advantage and 34 per cent. any disadvantage to day-release training. In considering the advantages, 14 per cent. said that there was more time for private study ; 11 per cent. said they appreciated the variety—the combination of practical and theoretical work ; and 10 per cent. mentioned the fact that they could be employed and earn more while studying. Of the disadvantages of day-release training, 19 per cent. of the students mentioned that there was not enough time and, therefore, the scope of study was limited ; 15 per cent. said that the day was too long and crammed ; this led to the student becoming too tired ; 13 per cent. said the courses tended to be disjointed because the time between lectures was too great.

Of the 174 students who had the experience of attending evening classes before starting the Diploma of Technology course, 28 per cent. had attended for more than 3 years, 47 per cent. from 1-3 years and 25 per cent. under 1 year. The majority of students attended once a week, although some students went two and three times. Again all students were asked to give their opinion on the advantages and disadvantages of evening training as compared with full-time study. 82 per cent. had no opinion on the advantages and 42 per cent. on the disadvantages of evening classes. Just as fewer students had considered the advantages of day-release schemes compared with the proportion who mentioned the disadvantages, so fewer students voiced any advantages of this way of

studying. Only 5 per cent. mentioned the advantage that the full practical training at work was uninterrupted while 35 per cent. mentioned that concentration was difficult after a full day's work when the student was tired. 12 per cent. mentioned that leisure time was curtailed by too many nights spent doing homework. 10 per cent. referred to the evening class syllabus being too condensed—there was not enough time for the subject—while 6 per cent. said progress was too slow ; it took too long to obtain a professional qualification.

*Entry Qualifications.* There were marked differences in entry qualifications for the Diploma of Technology between students. Basically, students were equipped in two ways—by school work leading to the General Certificate of Education ; by day release and evening class training for National Certificate examinations. It may also be added that the prior experience gained from working for a firm can be a valuable asset to a student embarking on the course.

TABLE 3

THE NUMBER OF SUBJECTS TAKEN BY STUDENTS AT THE GENERAL CERTIFICATE OF EDUCATION ORDINARY LEVEL.

Analysed by number of subjects passed at Advanced Level.

QUESTION : What examination qualification had you on leaving full-time school (specify subject and level) ?

No. of Subjects at G.C.E. 'O' Level	Total	No 'A' level subject	1 'A' level subject	2 'A' level subjects	3 'A' level subjects	4 'A' level subjects
Base of %—No. of students who had 'O' level qualifications .....	328	112	13	84	111	8
Proportion who passed :	%	%	Nos.	%	%	Nos.
One to three subjects ..	7	11	—	4	5	1
Four subjects .....	9	13	3	7	4	—
Five subjects .....	18	19	1	14	23	—
Six subjects .....	23	29	2	25	19	2
Seven subjects .....	20	15	4	22	23	1
Eight subjects .....	14	10	2	15	14	4
Nine subjects .....	6	2	1	8	7	—
Ten subjects .....	3	1	—	5	4	—
Eleven subjects .....	*	—	—	—	1	—

84 per cent. of the 392 students interviewed had 'O' level qualifications. Of these 328 students 88 per cent. had passed in Maths. ; 84 per cent. English Language ; 69 per cent. Physics ; 59 per cent. Chemistry ; 48 per cent. Geography ; 47 per cent. English Literature. 55 per cent. of students had 'A' level qualifications. Of these 216 students, 90 per cent. passed in Physics, 77 per cent. Mathematics and 36 per cent. Chemistry. Of the 158 grammar school and thirty-four public school students who had 'A' level qualifications, 59 per cent. and 50 per cent., respectively, had passed in three or more subjects. 112 Electrical Engineering students had 'A' level qualifications compared with seventy-nine Mechanical Engineering students. There is no significant difference in the proportions of students with 'A' level qualifications over the four years interviewed, nor between work and College based students.



Only 14 per cent. of students had come straight from full-time day school to the College of Advanced Technology. During the period between leaving school and attending the College, 52 per cent. gained other academic qualifications. Of these 204 students, 77 per cent. passed Ordinary National Certificate, a markedly higher proportion than gained any other qualification. During this period also many students gained experience that could prove a valuable preparation to the course; 62 per cent. had held one job; 16 per cent. two jobs; 8 per cent. more than two jobs. Of the 337 students who had had jobs, 62 per cent. had never taken a temporary job (one of three months or less) while 30 per cent. had had one temporary job and 8 per cent. two or more.

Because there are various means of entry upon a Diploma of Technology course, it is difficult to assess the academic attainment and experience upon which students are selected for that qualification—'the equivalent of an honours degree.'<sup>2</sup> Whether we can see in this experimentation in the expansion of higher education evidence for the confidence expressed in the Robbins' Report—"At this stage in the history of British higher education, it is a mistake to regard the claims of quantity and quality as being in conflict. Regard for the former is a safeguard against waste of talent; regard for the latter is a guarantee of the worth and merit of the whole"<sup>4</sup>—is doubtful. This survey seems to give further evidence to the student survey cited in the Robbins' Report that showed that "science not only attracts greater numbers but also attracts students of better quality than does technology."<sup>5</sup>

*What the students felt their education lacked as a preparation for the Diploma of Technology course.* In the light of their entry qualifications, it is worth considering what students themselves said their education lacked. A higher proportion of the sixty-two technical and thirty-two secondary modern students mention their inadequate knowledge of subjects than scholars from other schools; 19 per cent. and 25 per cent., respectively, mentioned Chemistry; 11 per cent. and 13 per cent., respectively, Mathematics; 11 per cent. and 50 per cent., respectively, English; 15 per cent. and 16 per cent., respectively, Social Studies and 10 per cent. and 22 per cent., respectively, Physics. Chemistry 15 per cent., Mathematics 15 per cent., English 15 per cent., Social Studies 10 per cent., and Physics 13 per cent., are mentioned by the 176 students who had no 'A' level passes. Apart from Social Studies, these subjects are mentioned by a higher proportion of 1st year students than those in their 2nd, 3rd and 4th year. The higher proportion of 2nd year students who become aware of their deficiency of knowledge in Social Studies may be accounted for by the compulsory English examination which must be passed before the end of the 2nd year. Of the 242 grammar and forty-two public school students, 11 per cent. and 10 per cent., respectively, drew attention to their lack of preparation in Engineering Drawing. Engineering Drawing is not cited by any students without 'A' level subjects in the General Certificate of Education. 44 per cent. of the 242 grammar school students, 32 per cent. of the sixty-two technical school students, 50 per cent. of the forty-two public school students and 19 per cent. of the thirty-two secondary modern students said their education lacked nothing.

*The Age students started the Diploma of Technology course.* On starting the Diploma of Technology course, 77 per cent. were under the age of 20 years, only 1 per cent. over the age of 23 years. 50 per cent. of the 355 sponsored students worked less than 1 year for their firms before being sent to the College. There were no significant differences in the intake age of the various years, departments, or between work and College based students.



## VI.—THE REASONS GIVEN BY STUDENTS FOR THEIR DECISION TO DO A DIPLOMA OF TECHNOLOGY.

The reasons given by the students for their decision to do the Diploma of Technology course fall into two main groups: (a) those referring to the qualification itself, and (b) those related to the advantages of the industrial experience. Of the first group of replies 17 per cent. mentioned the high qualification of the Diploma of Technology, while 20 per cent. said it was a second best to the university degree course for which they failed to meet the entry requirements. This latter point is elaborated on by Dr. Marie Jahoda. The Robbins' Committee offers one explanation in that "these Colleges are kept in a position of tutelage so that they are less attractive to students."<sup>6</sup> 13 per cent. mentioned that the Diploma of Technology was a full-time course preferable to day-release, night school or studying for an external degree. Of the second group of replies, 24 per cent. said that they chose to do the Diploma of Technology because of the practical training given with the theory; 13 per cent. mentioned their financial independence. Several students pointed out that by receiving a salary they were better off than many university undergraduates on a grant, which often had to be supplemented by parents or temporary holiday employment. 7 per cent. said they chose to do the Diploma because it was a useful training for an engineer, while 13 per cent. thought that it offered good prospects and they wanted to get on. Of the remaining replies, 8 per cent. said they had had no choice—either their firm had sent them or the College had transferred them from another course to do the Diploma of Technology. This was particularly the case of 3rd and 4th year students.

By far the most important people to influence students to study for a Diploma of Technology were their fathers (37 per cent.) and their firm's Educational Officer (34 per cent.). Only 18 per cent. however, admitted that their fathers had suggested specific jobs for which they should aim and of these seventy students 67 per cent. said their fathers had suggested engineering jobs. The proportion of fathers who suggested jobs for their children tended to decrease with lower social status.

All students were asked what they had considered to be the main advantages to be gained from doing a Diploma of Technology before they started the course and what they now considered to be the main advantages. From the similarity of the answers given, it seems clear that students did not remember what they had previously considered to be the main advantages of the course. Dr. Marie Jahoda makes a similar point in reference to her own work: "It is impossible to disentangle from the available material those attitudes with which the student embarked on his job from those which are the result of this experience."<sup>7</sup> The two most important advantages said to have been considered prior to embarkation on the course were the industrial experience combined with the theoretical study (43 per cent.) and the exemption from the examinations of professional bodies entailed by the Diploma of Technology (21 per cent.). More than half the Building and Civil Engineering students mentioned the former advantage.

In the light of their experience of the course, the industrial experience combined with theoretical study is mentioned by the highest proportion of students (44 per cent.) as one of its main advantages. The good education—the complete training of the course—receives more emphasis, 16 per cent. after, against 4 per cent. before, starting the course. 16 per cent. now mention the good qualification of the Diploma of Technology which gives exemption from examinations of professional bodies. These three advantages are considered by students of all departments and all years.



# VII.—STUDENT COMPARISON OF THE B.SC. TECHNOLOGY WITH THE DIPLOMA OF TECHNOLOGY.

The students were asked if they had considered doing a technological degree. 59 per cent. of the students said they had—83 per cent. of the forty-two public school students, 61 per cent. of the 242 grammar school scholars, 44 per cent. of the sixty two technical school pupils and 41 per cent. of the thirty-two secondary modern students. Of these 232 students who had considered doing the degree course, 51 per cent. said they decided against it because of lack of qualification; 24 per cent. for financial reasons; 18 per cent. because they wanted the industrial experience provided by a sandwich course; and 7 per cent. mentioned that they thought the Diploma of Technology course was better.

The remaining 41 per cent. of students said they did not consider doing a technological degree. Of this total of 160 students, 38 per cent. mentioned their lack of qualification, 13 per cent. their lack of money and 18 per cent. gave the positive reason that they wanted the industrial experience offered by the sandwich course.

When all students were asked what they considered to be the main advantages of doing a degree course, 31 per cent. said there was no advantage. The main advantages given were that the university offered a better social life (21 per cent.); the degree carried a prestige or snob value (16 per cent.); it was a recognised qualification, the Diploma of Technology was as yet not established (16 per cent.); the university course allowed more time for study as it extended over the whole year (13 per cent.). Only 2 per cent., however, thought that the degree course gave better prospects and led to higher wages.

Ideas that had already found expression were repeated when students were asked what they considered to be the essential similarities and differences between a Diploma of Technology and a B.Sc. degree in Technology. Concerning the similarities, 35 per cent. mentioned that both qualifications gave exemption from the examinations of professional bodies and, therefore, were of the same academic standard. 33 per cent. said the syllabuses covered the same subjects. A quarter of the students did not know in what ways the qualifications were similar. The most important difference mentioned was that the Diploma of Technology offered more industrial experience (55 per cent.). 13 per cent. thought that the degree had more theoretical training. There is disagreement between students as to the social status conferred by these two qualifications. 8 per cent. said that both led to similar jobs and salaries and, therefore, carried similar social status. 19 per cent. said that the essential difference between these qualifications was the 'snob value' attached to a degree. There seem to be four reasons why the Diploma of Technology does not have 'snob value.' The students believe that the Diploma of Technology is still not widely recognised as a professional qualification equal to an honours degree; this College of Advanced Technology has no long-standing traditions; students from every walk of life attend the course; many see their future work as practical rather than administrative or 'white collar.' The Diploma of Technology carries none of the advantages of the 'old school tie.'

# VIII.—DOUBTS EXPRESSED BY STUDENTS ABOUT THEIR CHOICE OF CAREER.

The students were asked whether they had any doubts about their choice of career. 77 per cent. said not. Of the remaining ninety-two students, 15 per cent. said they might have preferred a career in the services; 7 per cent. would have preferred to do a pure science degree and 5 per cent. another technology; 7 per cent. simply stated that they would have preferred work that was 'non-technical.' This latter proportion, taken with the 14 per cent. who

did not know what they would have preferred to do, shows a small nucleus of students who are discontented with the course without having a clear conception of an acceptable alternative training.

#### IX.—STUDENT PROSPECTS.

There is no clear picture of what students visualise will be their future in industry. Comparing their prospects with their father's present achievement, 51 per cent. thought they would be in a socially better position and 41 per cent. better off financially at the end of their course than were their fathers; 34 per cent. and 48 per cent., respectively, thought not. At the height of their career 66 per cent. thought they would have risen socially and 79 per cent. would have a better paid job; 10 per cent. and 7 per cent., respectively, were less hopeful. Of the eighty-four students from AB class homes, 17 per cent. thought they would, 74 per cent. would not, be socially in a better position than their fathers after training while 5 per cent. thought they would and 86 per cent. would not be financially better off. In contrast, 88 per cent. of the sixty-seven students from D<sub>2</sub>E class homes though they would have risen socially compared with 1 per cent. who thought not; while 79 per cent. thought they would, 18 per cent. would not be financially better off. Not all students were prepared to answer each section of this question.

When asked to state the position students would like to hold when they leave the College, 35 per cent. said they did not know. 14 per cent. said they wanted a position in administration without giving specifications and 5 per cent. wished to do research work—a few students from each of the firms analysed.

TABLE 4

THE POSITIONS STUDENTS WOULD LIKE TO HOLD WHEN THEY LEAVE COLLEGE.  
Analysed by Department.

QUESTION: State the position you would like to hold when you leave this College.

Preferred Future Positions	Total	Electrical Engineers	Mechanical Engineers	Building and Civil Engineers	Chemical Engineers
Base of %—All Students	392	173	173	27	19
Proportion who said :	%	%	%	%	Numbers
Position in Administration/Management . . .	14	15	13	7	6
Design and Development Engineer . . . . .	6	4	9	—	—
Research Worker . . . . .	5	5	7	—	—
Electrical Engineer . . . . .	4	9	1	—	—
Site Engineer . . . . .	3	1	1	—	—
Teacher/Lecturer . . . . .	3	3	3	30	—
Electronics Engineer . . . . .	2	4	1	—	—
Production Engineer . . . . .	2	1	3	—	—
Manufacturing Engineer . . . . .	1	1	—	—	1
Erection Engineer . . . . .	1	1	1	—	—
Laboratory Technician . . . . .	1	2	—	—	—
Section Leader . . . . .	1	1	—	—	—
Consulting Civil Engineer . . . . .	1	—	—	7	—
Outside Tech. Consultant . . . . .	1	1	1	4	—
Instrument Engineer . . . . .	1	1	1	—	—
Experimental Officer . . . . .	*	1	—	—	—
Other Engineering Posts . . . . .	13	10	13	15	5
Don't Know . . . . .	35	31	43	33	5
Other Replies . . . . .	6	9	3	4	2



Dr. Jahoda mentioned the lack of ambition of some of the students that she interviewed. From this College of Advanced Technology, 46 per cent. of students thought their chances of reaching the top of their profession were good or very good. 48 per cent. rated their chances less highly. Less than half the students (43 per cent.) said they knew a colleague who was likely to reach the top of his profession. Of these 170 students, 49 per cent. attributed the chances of his success to intelligence, ability, 29 per cent. to initiative, 25 per cent. to hard work, 25 per cent. to having a pleasant disposition. The fact that 4 per cent. mentioned that a colleague had a good background indicates some discrepancy between students in this respect.

72 per cent. of students said 'yes' when asked if they would be willing to pursue a further course of study when their Diploma of Technology course had finished. This desire seemed unaffected by any difference of age of the student on starting his Diploma of Technology course and was shared by similar proportions of students in each department. The proportion of students who would study further but did not know what subject to choose (16 per cent. of the total) tended to decrease from 1st to 4th year. Just under half the students, thought specialization was, or probably was, necessary to reach the top of their profession. During their first year at College (147) a greater proportion of students (26 per cent.) did not know whether further specialization would be necessary.

An investigation was made of the salaries students hoped to get in the future. Immediately on finishing training, 32 per cent. said they expected under £700 per annum; 31 per cent. between £701—£800; 19 per cent. between £801—£1,000 and 2 per cent. over £1,000. Within each social class the proportion of students gradually increases up the salary scale until a peak proportion is reached with a salary expectancy of between £701—£800; above this peak, the proportion decreases. As students get older a greater proportion expect a higher salary at the end of their course. Of the 173 students under 19 years, 39 per cent. expect under £700, 43 per cent. over £700; of the 196 19—21-year-old 28 per cent. expect under £700 and 57 per cent. over £700; of the twenty-three students 21 years and over, four expect under £700 and sixteen over £700. The salary expectancy is also greater among Mechanical than Electrical Engineering students. 20 per cent. of the 173 Mechanical Engineering students compared with 44 per cent. of the 173 Electrical Engineering students expect under £700; 58 per cent. Mechanical Engineering compared with 45 per cent. Electrical Engineering students expect over £700. £701—£800 is the salary expected by the largest number of students of any social class, age or department.

Students were asked what salary they hoped to get 5 years after their course had finished. 5 per cent. said under £1,000; 63 per cent. between £1,000—£1,500; 13 per cent. over £1,500. The social class and age of the students did not appreciably affect their salary expectancy. A lower proportion of Mechanical Engineering students as compared with Electrical Engineering students expected the lower salary ranges and a greater proportion expected the higher salary ranges.

When asked to estimate the salary they hoped to get at the height of their career, 16 per cent. of students said under £2,000, 39 per cent. between £2,000—£3,000 and 12 per cent. over £3,000. 54 per cent. of the 173 Electrical Engineering students compared with 44 per cent. of the Mechanical Engineering students expected salaries of over £2,000—a reversal on the answers given to the two previous questions although it must be born in mind that 38 per cent. of Mechanical compared with 29 per cent. of Electrical Engineering students



did not give any estimation of what salary they hoped for at the height of their career. 50 per cent. of both groups of students under 19 years and 19-20 years expected salaries of over £2,000.

To the three previous questions on what salaries students hope to get. 12 per cent. declined to answer the first; 14 per cent. the second and 17 per cent. the third. In each case less students declined from AB class homes than others; more students under 19 years declined than 19-20 years; and more Mechanical than Electrical Engineering students.

All students were asked to show by the numbers one to five the order of importance of five reasons for salary increase. An equal proportion of students gave priority to "because you make decisions affecting others in the firm" and "because you have increased technical skill." Most students thought age was the least important reason for salary increase.

#### X.—STUDENTS' PRESENT SALARY OR GRANT.

The questions on the students' present salary or grant were left to the end of the questionnaire in the hope that as many students as possible would answer. They were placed after those relating to what the students hoped would be their future salary, as it was felt that students on a small salary or grant at present might be more readily encouraged to disclose their income and whether they were supported by their parents, if it was clear that the Diploma of Technology course at the College was looked upon as a purely interim period and no guide to their earning capacity. 8 per cent. of the 311 students earning an annual salary did not say how much they earned and 1 per cent. refused to say whether or not their fathers helped to support them.

The wide range of salaries earned by students tended to lead to ill-feeling. Several College based students were envious of students earning a salary which often had increments attached for each year's progress. 79 per cent. of students received an annual salary—a higher proportion (86 per cent. to 76 per cent.) of the 173 under 19 years, than the 196 19-20 years. Of the 311 students who receive an annual salary 48 per cent. receive under £400, 25 per cent. £400—under £500; 19 per cent. £500 and over. Although covering the same range of salaries the proportion of students earning more rises with age. The majority of work based students who did not receive an annual salary came from the smaller sponsoring firms.

24 per cent. of all students said they received a grant. Of these ninety-three students, 19 per cent. said the grant supplemented a salary. 18 per cent. received under £50; 28 per cent. £51—£120 and 54 per cent. over £100 per annum. Those students who receive a grant that does not supplement a salary, were asked how much they earned during their six months' industrial training period. 1 per cent. said under £100; 41 per cent. £100—£199; 25 per cent. £200—£250 and 10 per cent. over £250. Those students who are older and receive a grant tend to earn more.

Father's financial help was received by 35 per cent. of all students—a slightly higher proportion from AB class homes than others, but a similar proportion for each age group.

I asked students what they thought their total annual income should be while a student. 6 per cent. said under £350; 35 per cent. £350—£499; 36 per cent. £500—£650; 5 per cent. said over £650. Of the 320 students who gave an estimate of what their salary should be, 72 per cent. thought they should receive more than at present; 22 per cent. were satisfied; 1 per cent. said they should receive less. As 5 per cent. had not given their present salary, it was impossible to ascertain whether their estimate of what they should receive was more or less than their present salary.



## XI.—CONCLUSION.

This extract is taken from a survey of Diploma of Technology students at a College of Advanced Technology, who will be among the first to pioneer this qualification in industry. Although the students are largely of local origin, they are drawn from a diverse social and educational background. For them, the Diploma of Technology is the opening door on a career formerly reserved for the graduate trainee. Recommendation 56 of the Robbins' Report that "these Colleges should, in general, be designated as technological universities, with power to award both first and higher degrees"<sup>8</sup> will do much to bring universal recognition and status to a sandwich course qualification. The outcome of this period of rapid expansion in advanced technological education will be assessed not only on the quality of the students but also upon whether the co-operation between the industrial firms concerned and the particular College of Advanced Technology is close enough to ensure that the sandwich course is an adequate preparation for the students available, for a responsible career in this changing technological age.

## XII.—REFERENCES.

Higher Education Report under the chairmanship of Lord Robbins (1963) :

<sup>1</sup> Ch. 10, page 146, para. 444, conclusion.

<sup>2</sup> Ch. 10, page 141, para. 430.

<sup>3</sup> Ch. 10, page 138, para. 417.

<sup>4</sup> Ch. 19, page 266, para. 835.

<sup>5</sup> Ch. 10, page 126, para. 378.

<sup>6</sup> Ch. 10, page 131, para. 391.

Nuffield Research Project Progress Report 1, by Dr. Marie Jahoda (1960) :

<sup>7</sup> page 17.

<sup>8</sup> Recommendation 56, page 281.

*(Manuscript received 18th December, 1963)*

# THE ATTAINMENTS IN ENGLISH AND ARITHMETIC OF SECONDARY SCHOOL PUPILS WITH IMPAIRED HEARING

BY D. C. WOLLMAN

*(Department of Audiology and Education of the Deaf, University of Manchester)*

**SUMMARY.** This paper reports the findings of an investigation into the attainments in English and arithmetic of a representative sample of 14-16-year old pupils with impaired hearing and considers factors which appear to influence such attainments.

## I.—INTRODUCTION.

THIS study formed part of a survey of educational provision, curricula and methods of teaching for secondary pupils with impaired hearing (Wollman, 1961). Its aim was to explore a possible basis for an examination which would serve as a leaving certificate and as a means of assessing the comparative attainments in English and arithmetic of 14-16-year old pupils with impaired hearing.

## II.—PLAN OF THE INVESTIGATION.

It was difficult to find standardised English tests for pupils with impaired hearing who are likely to be retarded linguistically. Two tests were, therefore, specially designed. The first consisted of seventy items, including correct and incorrect sentences, prepositions, colloquial expressions, spelling, completion of compound and complex sentences and vocabulary. The second was a composition test in which pupils were asked to write three short passages:

- (a) Describing some object or experience, e.g., my home, my favourite television programme.
- (b) Explaining some activity, e.g., lighting a fire, cleaning shoes.
- (c) Composing a formal letter, e.g., applying for a job.

A choice of subject was allowed for each passage.

Two arithmetic tests were used. The Manchester Mechanical Arithmetic Test was considered specially suitable as it covered a wide range of processes including area, decimals and percentages which normally figure in the curriculum for older pupils. In the second test an attempt was made to provide arithmetical problems relating to familiar experiences.

In the first English test and the two arithmetic tests, one mark was given for each correct item. In the composition test a marking schedule was based on a five letter scale (A—E) in which the mark C was awarded if the material was intelligible and relevant even though there were mistakes in language structure. Using this schedule all papers were marked by the writer and two colleagues experienced in the assessment of the language development of pupils with impaired hearing. The product moment correlations between markers were of the order of 0.9. The final mark awarded to each piece of work was the average of the marks given by each examiner.

The lower age limit of the experimental group was set at 14 years 0 months at the time of the test and the upper limit at 16+. It was not practicable to test the whole population of this age range. The aim was that the experimental group should be reasonably large and drawn from a representative selection of schools for children with different degrees of impaired hearing. The head teacher



and staffs of thirteen schools for the deaf and/or partially hearing agreed to co-operate. The total number of pupils tested represented between a quarter and a third of pupils in these age groups being educated in special schools (Min. of Education, 1959).

Hearing losses were assessed by finding the average of the hearing loss in the better ear in decibels (db) above normal threshold at the frequencies 500, 1,000 and 2,000 cycles per second. Pupils were divided on the basis of average hearing loss into two groups with a dividing line at 75 db. This level was chosen to separate those children who hear unamplified speech albeit imperfectly and those with more severe impairments who, in the main, have little or no experience of unamplified speech.

The two English tests were also given to 162 pupils (ninety-five boys and sixty-seven girls) at two secondary modern schools. Their ages ranged from 14 years 3 months to 15 years 6 months, with a mean of 14 years 9 months.

### III.—THE FINDINGS OF THE INVESTIGATION.

(1) In both English and arithmetic the pupils with impaired hearing were found to be retarded when compared with unhandicapped pupils. Their mean scores in the English tests were significantly lower than those of the secondary modern pupils. In the standardised arithmetic test they had a mean raw score of 14; this would represent a standard score of 85 for boys and 87 for girls (norm. 100, S.D. 15).

(2) Amongst the deaf pupils, those with moderate impairments of hearing tended to obtain higher scores in English. Degree of hearing impairment did not appear to be a factor in performance in arithmetic.

(3) Mean English scores for girls were higher than those for boys; but only amongst the severely deaf pupils was the difference large enough to be statistically significant. The superiority of the severely deaf girls was particularly marked in the composition test. Sex difference was not apparent in the results of the arithmetic tests.

(4) The mean scores for the 14 and 15-year age groups were approximately equal. In other words, there was no evidence of any increment of score through increase in age. The two age groups did not show variations in distribution of intelligence or hearing loss.

(5) In the first English test the greatest success was achieved in those items where the emphasis was on comprehension rather than expression. Many pupils were successful with the questions about a prose passage which were worded in a straight-forward manner but had difficulty with those in which the language structure was more involved. The pupils with impaired hearing scored as well or better than the secondary modern pupils in the abbreviations, plural forms and spellings where reliance on a visual pattern may tend to produce greater accuracy.

In items requiring expression many pupils, especially the severely deaf, understood the idea required, but were not familiar with the conventional method of expression, e.g. 'The ball went *break* the window.' This was clearly demonstrated in those items where sentences had to be completed so as to make them intelligible and grammatically correct. All pupils, hearing and deaf, found sentences containing 'across which' and 'although' the most difficult and those with 'and,' 'but' and 'because,' the easiest. The item with 'although' was successfully completed by 39 per cent. of secondary modern pupils, 12 per cent. of partially hearing pupils and 4 per cent. of the severely deaf; for the sentence containing 'but' the comparable figures were 70 per cent., 48 per cent. and 21 per cent. The differences between the percentages of

pupils passing in the three groups gave some indication of the effect of hearing impairment on the understanding and expression of complex sentence patterns.

An analysis of answers in this section suggests that the difficulties encountered were of two main types: (i) The sense of some conjunctions was not understood. Clauses of cause instead of concession were supplied after 'although' and in many instances the word 'because' was inserted. The use of a preposition with a relative pronoun caused confusion to many pupils, including a number from the secondary modern schools. (ii) Conventional language patterns were not known. The most frequent errors were: (a) the use of a single word to express an idea (The child was knocked down while *a car*). (b) errors in grammar, such as non-agreement of subject and verb, incorrect participles and wrong tenses.

(6) The compositions in the second English test were analysed to discover the average length of composition, average length of sentence and types of sentence used. The figures for length of composition are shown in Table 1.

TABLE 1

AVERAGE NUMBER OF WORDS PER COMPOSITION FOR BOYS AND GIRLS IN EACH HEARING LOSS GROUP.

Question No.	Hearing Loss greater than 75 db.		Hearing Loss less than 75 db.		Secondary Modern	
	Boys	Girls	Boys	Girls	Boys	Girls
1	103.1	133.8	92.2	144.6	116.0	126.3
2	57.6	68.7	64.1	74.0	84.0	87.2
3	41.3	47.9	52.9	60.4	51.7	61.2

In all questions and amongst all pupils the average length of the girls' compositions was greater than that of the boys'. The tendency was for the deaf pupils to produce shorter compositions.

On the whole, the sentences were clearly enough marked off to make the count reliable and the classification straight-forward, but there were instances of incomplete or unclassifiable expressions which had to be excluded from the analysis. The figures for length of sentence are given in Table 2.

TABLE 2

THE AVERAGE LENGTH OF SENTENCES FOR BOYS AND GIRLS IN EACH HEARING LOSS GROUP.

Question No.	Hearing loss greater than 75 db.		Hearing loss less than 75 db.		Secondary Modern Pupils	
	Boys	Girls	Boys	Girls	Boys	Girls
	Average number of words per sentence					
1	8.6	9.4	11.1	11.3	14.0	13.8
2	8.87	9.8	11.57	11.84	14.4	14.1
3	8.83	10.42	11.8	11.8	15.8	16.2



Variation in average length of sentence was seen to be related to degree of hearing impairment. Differences in sentence length between boys and girls were only apparent amongst severely deaf pupils where the girls' average was rather higher.

Four classes of sentences were distinguished—simple, compound, complex and compound-complex. Table 3 shows the proportion of different types of sentences used by pupils.

TABLE 3  
PROPORTION OF DIFFERENT TYPES OF SENTENCES USED.

Type of Sentence	Hearing loss greater than 75 db.	Hearing loss less than 75 db.	Secondary Modern Pupils
	Percentage of each type of sentence		
Simple .....	74	56	34
Compound .....	12	15	19
Complex .....	11	21	31
Compound-complex ..	3	8	16

The analysis supported the finding in the first English test that degree of hearing impairment was an important factor in the complexity of written expression. The pupils with impaired hearing used twice as many simple sentences as the secondary modern pupils. Three-quarters of all sentences written by severely deaf pupils were simple sentences; approximately half the sentences produced by partially hearing pupils and a third of those produced by the secondary modern pupils were of this type. Pupils with losses less than 75 db used twice as many complex and compound-complex sentences as those in the other hearing loss group.

Enumeration was found to be a feature of many compositions produced by severely deaf pupils. It occurred even in some of the better ones, e.g., Girl, 16.4, 100 db. "My home have four windows and a door at the front of a house, three windows and a door at the right side, five windows at the back." In the worst there was merely a list of nouns or a string of echoic sentence patterns, e.g., Boy, 14.0, 105 db. "I have television, table, cat, bicycle, chair, flower and stool."

There was a tendency amongst the severely deaf pupils to describe or narrate in minute detail. In some instances, when beginning to write about "A day in my holiday," pupils would explain that they got up, washed, dressed, went downstairs, had breakfast, got in the car—a series of minor details which hearing pupils in the age range would take for granted.

Although some writers did not follow a recognisable pattern of English, it was usually possible, with a back-ground of language experience, to interpret what the writers were trying to say. In some compositions written by severely deaf pupils confusion of word order suggested translation from a non-syntactical sign language. Boy, 15.0, 85 db. "My bicycle from Christmas home at out play time pay shops £18 19s. 6d. 36 week. Father by pay bicycle shops."

Errors in grammar and construction were found in the compositions of all pupils but were more common amongst the severely deaf. Articles and prepositions were omitted or were included when they were not required. In some sentences the wrong preposition was used. It was apparent that the majority of pupils understood the function and were able to use successfully, the present, past and future tenses, but there were many instances where incorrect forms of verbs were used. Irregular verbs, participles and the infinitive were frequent sources of error.

Almost half the pupils with impaired hearing were able to compose a letter which was regarded as adequate. Of the others, the main criticisms would be social immaturity and lack of experience (practical or verbal) of the situations with which the letters dealt.

(7) A large proportion of all the pupils showed a skill in the simple arithmetical processes (four rules); the majority gave no indication of ability in more advanced work. Over half showed little evidence of knowledge of reduction and of tables of weight, length, etc.; between two-thirds and three-quarters were apparently unable to deal with fractions, decimals and the more difficult processes. There was little variation between the performance of pupils in the two hearing loss groups but such difference as existed was in favour of the severely deaf pupils.

The majority of those tested were able to deal successfully with simple arithmetical problems. Far more errors were attributable to lack of understanding than to mistakes in calculation. In a problem about the cost of fencing a plot of land, the length of one side was given as the answer or was multiplied by two or three; some pupils multiplied the length by the breadth, showing confusion between this process and area.

(8) Variation in test performance amongst schools was investigated. The results support K. P. Murphy's findings (Murphy, 1956) of a wide scatter of test scores from school to school and his conclusion that these differences in attainment cannot be explained entirely by variations in intelligence or degree of hearing impairment. There was little variation in performance between day and resident pupils. Those with moderate impairments of hearing who were resident in schools catering for both deaf and partially hearing children tended to obtain lower scores in English.

#### IV.—DISCUSSION AND CONCLUSIONS.

(1) The marked difference in English attainments between severely deaf girls and boys is confirmed by Myklebust (1960). The explanation may be that, for the handicapped, the psychological changes of adolescence are delayed and, therefore, boys remain linguistically inferior to girls until a later age.

(2) The similarity of performance of the 14 and 15-year age groups may be partly accounted for by the organisation of educational provision. Many schools are so small that it is not possible to arrange a system of streaming; brighter children may reach the top class at the age of 12-14 years and complete their school career there. As they are working alongside older pupils with less ability and poorer attainments, they may lack the stimulus of competition and may accept the standards of their classmates. Unless they have a teacher enterprising enough to plan progressive schemes of work to suit individual requirements and to keep pupils interested and working to capacity, there will be the danger of educational stagnation.



It seems unlikely that secondary education for pupils with impaired hearing will be fully implemented until it is organised in separate schools for primary and secondary age ranges. This would make possible staffing ratios and scales of equipment appropriate to the secondary stage of education. It would encourage the use of methods appropriate to adolescents which would give reality to the work done in school and thus provide motivation to higher standards of attainment. One way would be to provide specialist courses of the type used in many secondary modern schools.

(3) Results in arithmetic suggest that the elementary processes are receiving too much attention to the neglect of more advanced subjects which would be of greater interest to older pupils.\* Teachers are seeking to obtain perfection in fundamental processes before venturing further. Three points arise from this :

(a) The need for careful consideration about what is fundamental in arithmetic. A number of topics appear to owe their inclusion in the curriculum to the influence of tradition rather than to their intrinsic importance.

(b) Constant revision and consolidation of basic processes is likely to lead to frustration and lack of interest.

(c) Arithmetic at the secondary stage could be given a greater sense of purpose by being linked to other subjects of the curriculum.

(4) Educational environment is an important factor in the standards of achievement of all children ; it is of special consequence to those with impaired hearing. Normal children learn a great deal incidentally outside the school environment. Much of this incidental learning and especially that which is dependent on auditory clues is denied to the deaf child and the school has to make good the deficiency. It is to be expected, therefore, that differences between the quality of education provided will have greater effect in schools for hearing-impaired children. Some factors which may influence the quality of education are :

(a) Historically, most of the earlier schools were institutes for the deaf and dumb where the aim was to fit pupils to earn a living principally by semi-skilled forms of labour. Some schools have inherited accommodation and physical conditions which are far from satisfactory.

(b) It was partly as a result of private effort that secondary education, in the first place for selected pupils, began to be developed. More recently non-selective secondary schools have been established but at the time of the investigation most pupils were not receiving a specific secondary education.

(c) The wide scatter of results amongst schools may reflect the lack of generally accepted standards for pupils with impaired hearing. Their attainments have been measured against norms for the unhandicapped. This has tended to emphasise retardation rather than to focus attention on positive achievements. It is hoped that the analysis of the results of the tests in the present investigation will serve to make more widely known the attainments of such pupils.

(d) Recent research (Ewing 1957, 1960) has indicated the importance for pupils with impaired hearing of the provision of auditory experience in furthering linguistic development and general educational progress. At the time of this investigation there were considerable differences between schools both in the amount of electronic equipment available and in the efficiency with which such equipment was used.

\*Min. of Ed. 1958.

It is not suggested that attainments in English and arithmetic—or, indeed, in other subjects—are or ought to be the sole criteria of educational achievement. Modern thinking rightly stresses the importance of the school in providing for personality, social, physical and emotional development and for helping the handicapped to adjust to their disability ; but, in a literate society, those with hearing impairments need to achieve adequate standards of communication (both spoken and written) if they are to integrate fully into that society.

#### V.—REFERENCES.

- EWING, A. W. G. (Ed.) (1957). *Educational Guidance and the Deaf Child*. Manchester : Manchester University Press.
- EWING, A. W. G. (Ed.) (1960). *The Modern Educational Treatment of Deafness*. Manchester : Manchester University Press.
- MINISTRY OF EDUCATION (1958). *Pamphlet No. 36—Teaching Mathematics in Secondary Schools*. London : H.M.S.O.
- MINISTRY OF EDUCATION (1959). *Education in 1958*. London : H.M.S.O.
- MURPHY, K. P. (1956). *A Survey of the Intelligence and Abilities of Twelve-year-old Deaf Children*. Unpublished thesis. Manchester University.
- MYKLEBUST, H. R. (1960). *The Psychology of Deafness*. New York : Grune and Stratton.
- WOLLMAN, D. C. (1961). *Some Problems Involved in the Application of Secondary Modern Education for Deaf Pupils*. Unpublished thesis. Manchester University.

(Manuscript received, 29th January, 1964.)



# A FURTHER STUDY OF SOME COGNITIVE AND OTHER DISABILITIES IN BACKWARD READERS OF AVERAGE NON-VERBAL REASONING SCORES

By K. LOVELL, E. A. GRAY AND D. E. OLIVER

(*University of Leeds*)

**SUMMARY.** The application of the Watts-Vernon Reading Test and a Non-Verbal Reasoning Test to the whole of the fourth year pupils in fourteen secondary modern schools gave 426 pupils with a reading quotient  $<80$ ; of these 204 had standardised non-verbal reasoning scores  $\geq 90$ . Thirty backward boys and twenty-five backward girls, all of whom had non-verbal reasoning scores  $\geq 90$ , were paired, individually, with average to good readers, matched for non-verbal score, social class, sex and school. The groups undertook a number of individual tests. The boys who were backward had a poorer performance on certain tests compared with non-backward boys, and a greatly inferior performance to backward girls on the tests of copying and dictation.

## I.—INTRODUCTION AND PROBLEM.

In an earlier paper in this *Journal*, Lovell, *et al.* (1964) presented evidence suggesting that:

- (i) Among backward readers (R.Q.  $<80$ ) in the junior school children aged 8 and over, roughly half had average or better non-verbal reasoning scores.
- (ii) When backward readers in the junior school were matched with average to good readers for non-verbal test score, social class, sex and school the backward readers showed a poorer performance on a variety of tests involving spatial relationships and left right discrimination; also a greater 'rotation' effect on a test involving the copying of abstract designs.

The present study involved children in the 14-15-year-old age group. The Watts-Vernon Reading Test and the N.F.E.R. Non-Verbal Test No.3 were given to all the fourth year pupils in fourteen secondary modern schools. Data were obtained on 981 boys and 872 girls. It was found that 6.3 per cent. of the boys and 3.3 per cent. of the girls had a raw score on the reading test of  $<9$  (R.A.  $<9$  yr.) and 11.4 per cent. of the boys and 7.3 per cent. of the girls had a raw score of  $<12$  (R.A.  $<10$  yr.). Once again the data shows that gross reading backwardness is far more often found in boys than in girls. This is so in spite of the fact that over the whole range of reading attainment, the test gives a slight advantage to boys. Equally interesting was the fact that of 426 pupils with an R.Q.  $<80$ , no fewer than 204 or 47.8 per cent. had standardised non-verbal reasoning scores  $\geq 90$  (upper limit 131). Clearly the thesis that roughly half of all backward readers have average or better non-verbal reasoning skills is as true among school leavers as it is in the upper junior school.

In an attempt to probe further into the disabilities of the backward readers of average or better non-verbal reasoning scores, thirty such boys and twenty-five such girls were selected for detailed study. These pupils, who were drawn from the 14-15-year-old group tested before, were examined by means of the tests listed below. It was also hoped that further light would be thrown on the differences in abilities between backward boys and backward girls. Each of these backward readers (R.Q.  $<80$ ) was paired with an average to good reader, the members of each pair being matched for non-verbal reasoning score, social class, sex and age, while both members of each pair were pupils

of the same school. In the case of the girls the matching on the basis of social class, and taking into account the other variables, was difficult in five of the twenty-five instances. It must also be stressed that the backward readers were the 'run of the mill' backward pupils in secondary schools and not those to whom special labels had been attached.

## II.—PROCEDURE.

It is known that it becomes increasingly difficult to distinguish between good and poor readers as children get older, except by using attainment type tests. It was realised that several of the tests used in the earlier study (Lovell, *et al. op. cit.*) would fail to discriminate between the members of the matched pairs in the 14-15-year-old group, and the battery of tests finally decided upon is listed below. All tests were administered individually to each child in his own school.

(1) WISC Vocabulary test.

(2) WISC Block Designs test.

(3) Thurstone's revision of the Gottschaldt Figures (Thurstone, 1943). One mark was given for each design correctly identified, the maximum possible score being 40.

(4) Shapiro's test of Rotation (Shapiro, 1962). The 'rotation effect' refers to a tendency to reproduce abstract designs correctly, but in a mis-orientated position. Twelve designs were used and the test was carried out by the subjects drawing each design on the fresh page of a pad.

(5) Eight of the designs used by Bender (1938) or found in the Ellis test (Bender, *op. cit.*) were included together with a further three, more difficult, designs made up by the senior author (K.L.). The subject copied the eleven designs, using a fresh page of a pad for each drawing. The marking scheme was close to that laid down by Wood, *et al.* (1940). Two marks were given for each drawing, and one mark for an error or for two errors symmetrically consistent. Reversing, inverting or turning through 90° was counted as an error. This scheme would not quite cover all eventualities and a small addition to the scheme had to be made.

(6) The test of spatial orientation designed by Semmes, Weinstein, *et al.* (1955). This required the subject to follow on the ground, routes represented on maps. Five of the maps were perceived visually; nine others, which were rotated and mirror images of the visual maps, were perceived through touch alone. The maximum possible scores for the visually- and tactually-guided responses were 35 and 65, respectively.

(7) A sentence copying test (Lovell, *et al. op. cit.*). Ten sentences had to be copied by the subject. The sentences contained words of one letter, two letters, etc., up to ten letters, and there were five examples of each word of a given length. The number of errors made were recorded under the headings: reversal of letter, distortion of letter, indecipherable word (the number of errors being recorded as the number of letters in the word), letter or word added, letter or word omitted.

(8) Dictation test. The ten sentences already copied had later to be written as the experimenter dictated them. The scheme for scoring the errors was exactly as in (7) above.

## III.—THE RESULTS.

Table 1 shows the initial data on the various groups of subjects. The following Table shows the results obtained on the first six tests. In the case of the Shapiro test, the raw scores were subjected to a logarithmic transformation, because of their negative skewness, in an attempt to normalise them for the purpose of testing for a significant difference in mean scores between backward



and non-backward readers. In Table 3 data are given for backward and non-backward pupils in the tests of copying and dictation. The total number of errors are given under the various headings and the mean error rate in brackets.

TABLE 1  
INITIAL DATA RELATING TO THE MATCHED GROUPS OF CHILDREN.

	Boys 30 pairs		Girls 25 pairs	
	Backward Group	Non- Backward Group	Backward Group	Non- Backward Group
Mean C.A. ....	14.5	14.5	14.6	14.5
Mean Reading Score ....	9.70	25.03	11.24	25.68
Range of Reading Score ..	3—14	22—31	6—14	21—32
Mean Non-Verbal Test Score	100.5	99.7	103.4	102.8
Range of Non-Verbal Test Scores .....	90—123	90—119	91—131	90—127
Social Class .....	II 3 III 18 IV 6 V 3	II 3 III 18 IV 6 V 3	II 3 III 11 IV 8 V 3	II 1 III 16 IV 5 V 3

TABLE 2  
DATA OBTAINED ON TESTS 1-6 FOR MATCHED GROUPS.

	Boys (30 pairs)			Girls (25 pairs)		
	Backward Group	Non- Backward Group	Difference Significant at 1% level	Backward Group	Non- Backward Group	Difference Significant at 1% level
Scaled score on WISC vocabulary	7.67	11.06	Yes	6.92	10.36	Yes
Scaled Score on WISC Block Designs	12.17	12.30	No	12.36	12.0	No
Number of correct responses on Gottschaldt Figs.	788	847	Yes	657	678	No
Raw score on Shap- iro Rotation Test (Mean Rotation on Designs for each subject) ..	9.96°	9.32°	No	11.13°	8.88°	No
Points scored on Bender-type test	535	577	Yes	464	470	No
Number of correct responses on Wein- stein test :						
Visual .....	890	898	No	782	750	No
Tactual .....	1496	1525	No	1294	1241	No

TABLE 3

TOTAL NUMBER OF ERRORS AND MEAN ERROR RATE.

Pupils	Reversal of letter	Distortion of letter	Incorrect letter	Total No. of letters in indeci- pherable words	Letter or word added	Letter or word omitted
Boys' Copying						
Backward . . . . .	2 (.07)	15 (.50)	68 (2.27)	0	26 (.87)	8 (.27)
Non-Backward . .	1 (.03)	0	14 (.47)	0	8 (.27)	12 (.40)
Girls' Copying						
Backward . . . . .	1 (.04)	9 (.36)	19 (.76)	0	9 (.36)	3 (.12)
Non-Backward . .	0	3 (.12)	6 (.24)	9 (.36)	5 (.20)	10 (.40)
Boys' Dictation						
Backward . . . . .	48 (1.6)	17 (.57)	409 (13.63)	581 (19.37)	416 (13.87)	118 (3.93)
Non-Backward . .	7 (.23)	0	42 (1.40)	31 (1.03)	40 (1.33)	21 (.70)
Girls' Dictation						
Backward . . . . .	19 (.76)	4 (.16)	136 (5.44)	57 (2.28)	131 (5.24)	62 (2.48)
Non-Backward . .	0	0	18 (.72)	27 (1.08)	18 (.72)	19 (.76)

## IV.—DISCUSSION.

The findings may be summarised as follows :

(a) There was a significant difference between the backward and non-backward readers on the WISC Vocabulary test. This was true for both sexes. Note carefully, however, that the mean vocabulary score for backward boys was not lower than that of the backward girls (c.f. (ii) below).

(b) For neither sex was there any difference between the backward and non-backward pupils on WISC Block Designs, Shapiro's Rotation test, or on the test of spatial orientation devised by Semmes, Weinstein, *et. al.*

(c) The non-backward boys performed significantly better than the backward boys on the Bender-type test and on the Gottschaldt test of Hidden Figures. It seems likely that these differences reflect perceptual difficulties in the backward male reader which are more readily detected at lower ages. There were no differences on these tests in the case of the girls.

(d) On the test of sentence copying, the number of errors made under some headings was very small. But for both sexes the most frequently found errors among the backward children were : distortion of letter, incorrect letter, and letter or word added. For these three types of error the backward reader made more errors than his paired partner.

(e) The results of the dictation test were very revealing. There were a large number of errors under the headings : incorrect letter, number of indecipherable words, letter or word added, letter or word omitted. The backward readers made many more errors than their partners, but even more significant, was the far higher number of errors made by backward boys than by backward girls.



From the overall evidence presented in this paper, four items should be considered together :

(i) Marked reading failure remained about twice as frequent in the male as in the female even at 14-15 years of age.

(ii) In spite of the fact that the mean reading score of the backward boys was  $1\frac{1}{2}$  points lower than that of the backward girls, the mean WISC Vocabulary score of the backward boys was not below that of the backward girls.

(iii) There were still differences between backward and non-backward readers, in the case of the boys, on two tests involving visual perception. This was not so in the case of the girls.

(iv) The mean error rate for certain important kinds of error was higher for boys than for girls on the copying test, and very much higher on the dictation test. The  $1\frac{1}{2}$  points difference in mean reading scores between backward boys and girls could never explain the great difference in spelling between the groups in the dictation test.

In total these four items strongly suggest that marked reading disability in the 14-15-year-olds is often one symptom of a wider syndrome or syndromes which affect the male more severely than the female. The same seemed true at 9-10 years of age in the earlier study (Lovell, *et. al op. cit.*). Differences in reading and spelling may occur between backward boys and backward girls, when matched on other variables, although oral vocabulary may not be affected. Tests of visual perception do not easily discriminate between good and poor readers at this age, although Holt-Hansen (1963) claims to have produced a test of visual perception that will help in the diagnosis of 'word blind' adults. It is hoped that this will also help in studying the aetiology of marked reading failure in school children.

From the point of view of the teaching situation in the classroom, it must be stressed, again, that many backward children are 'underestimated.' Our data relating to the non-verbal reasoning scores show clearly that many backward pupils are capable of handling a wide range of ideas, of assimilating much information, and of a satisfactory level of performance in many skills, providing reading and writing do not make too great demands upon them. At present, these pupils are likely to be insufficiently exercised in many areas in the school situation.

#### V.—REFERENCES.

- BENDER, L. (1938). A visual motor Gestalt test and its clinical use. *Res. Monog. Amer. Ortho-psychiat. Assoc.*, No. 3.
- HOLT-HANSEN, K. (1963). *New Pathways in Understanding Symptoms of Word-Blindness based on Research in Visual Perception*. Paper read at the 17th International Congress of Psychology, Washington, 1963.
- LOVELL, K. *et al* (1964). A study of some cognitive and other disabilities in backward readers as assessed by a non-verbal test. *Brit. J. Educ. Psychol.*, 34, 58-64.
- SEMMES, T., WEINSTEIN, S., GHENT, L., TEUBER, H. L. (1955). Spatial orientation in man after cerebral lesion. *J. of Psychol.*, 39, 237-244.
- SHAPIRO, M. B., *et al.* (1962). Experimental studies of perceptual anomaly VII. *J. Ment. Sci.*, 108, 655-668.
- THURSTONE, L. L. (1943). *A Factorial Study of Perception*. Chicago: University of Chicago Press.
- WOOD, L., and SHULMAN, E. (1940). The Ellis visual designs test. *J. Educ. Psychol.*, 36, 591-602.

(Manuscript received 14th February, 1964)

# IMAGERY AND MENTAL PRACTICE

BY K. B. START\* AND A. RICHARDSON

(*University of Western Australia*)

**SUMMARY.** The vividness and autonomy of imagery of thirty-two males, aged 18-21 years, was compared with their performance scores obtained at their first attempt to perform a physical skill which they had been mentally practising for six days. First order correlations between vividness of Visual, Kinaesthetic and General Imagery, and the criterion performance scores were insignificant. Autonomy of imagery had a chi square with the criterion which fell marginally short of 5 per cent. significance, but this level was achieved using the Mann-Whitney U Test. When combinations of autonomy and vividness were compared with criterion scores it appeared that vivid autonomous imagery was associated with low criterion scores in this sample.

## I.—INTRODUCTION.

IN activities requiring complex co-ordinated physical movements such as in gymnastics and diving, competitors frequently use a mental practice (symbolic rehearsal) immediately prior to actually performing the physical skill. This habit is considered to be of considerable benefit to the subsequent skilled performance. Recently there have been many studies which have experimentally demonstrated the improvement in physical performance which can follow specific mental practice of the movement. It is now evident that a physical skill may be 'learned' without any overt physical practice being involved. It remains to be established whether this 'learning' is habit reinforcement at a subliminal level in the neuromuscular system as the studies in Electromyography (Davies<sup>5</sup>, Humphrey<sup>11</sup>, Jacobson<sup>12</sup>, Shaw<sup>16</sup>) and muscle tension (Freeman<sup>7</sup>, <sup>8</sup>, Williams<sup>24</sup>) might suggest, or whether it is a purely cortical phenomenon either one of 'grooving' the ability to concentrate on the task and so establish a favourable 'set,' or of establishing neurological pathways in the higher centres associated with proposed physical skill.

Investigations of the efficiency of mental practice have shown that it can vary with the physical familiarity of the skill to be rehearsed (Clark<sup>4</sup>), and even when the subject is naive to the motor skill, specific mental practice can be of value in improving physical performance (Start<sup>18</sup>, <sup>19</sup>). Pear<sup>15</sup> made the qualitative statement in 1924 that "some people who have never executed certain combinations of movements can occasionally compound them in imagination, making a passable performance when the time of action comes" (p. 44). This would suggest that efficiency of imagery might be an important factor in the ability of a person to improve in motor performance after mental practice, and this was the hypothesis which was investigated in the present study.

## II.—TESTS.

Imagery varies in vividness, modality and autonomy. Tests of these three dimensions rely on introspection and because of the subjective nature of this information, it is without verifiability by a second observer. From Galton's<sup>9</sup> early investigations into the nature of imagery, Betts<sup>2</sup> drew up an imagery inventory. Sutcliffe<sup>20</sup> produced an Imagery Battery derived from Bett's

\* Now school of Education, University of Manchester, England.



inventory and this was used to measure modality and vividness in the present study. In this test the subjects were asked to evaluate on a seven-point scale the vividness of the image developed in response to five cue sentences on each of seven modalities, visual, auditory, touch, kinaesthetic, taste, smell and organic.

The ability of an individual to manipulate his imagery is now regarded as being of consequence in the formation of stereotypes (Gordon<sup>10</sup>, McKellar<sup>12</sup>) and in this study it was felt that a person who had a static imagery would be unable to adequately picture the sequence of movements required for the criterion motor skill. The Gordon Test<sup>10</sup> for the autonomy of imagery was used to establish if the subject could alter his images or "if the first image they have is apt to hold its place tenaciously" (<sup>9</sup>, page 75). Basically, the test required the subject to image a particular situation and then vary it at the instruction of the tester, e.g., "Picture a red car—now can you see it as a blue car?—now image someone walking towards it—getting into it—driving it away." The ease with which the subjects could make each of these changes was taken to be a measure of the control they had over their imagery.

The criterion skill selected was the gymnastic movement known as Single Leg Upstart on the High Bar. The movement is regarded as gymnastically simple and pilot studies had indicated that it would provide a discriminating mental practice criterion. The skill required specific equipment which was in the control of the investigator and alternative apparatus was not readily available. In these circumstances, physical practice was regarded as being unlikely during the period of the experiment.

To standardise the mental practice the skill was analysed and the important movements recorded in sequence. A duplicated instruction sheet was compiled and this gave the logical analysis of the movement as well as instructions at various stages of the analysis for the subjects to imagine the movements described earlier. A modified form of this instruction sheet was used by the tester to enable him to maintain uniformity during the practice periods. All subjects were given six daily five-minute periods of this standardised mental practice of the gymnastic skill and on the seventh day they were asked to perform the criterion movement for the first time. In the following week the subjects were given the vividness of imagery test and during the second week they were given the test for the autonomy of their imagery.

The initial sample consisted of fifty-two males between the ages of 18-21, all of whom were in their first year at a secondary teachers' college in Australia, and were naive to the criterion skill. Forty-four subjects completed the requirements of full attendance at the mental practices and subsequently took the tests of physical performance and of vividness of imagery. Unfortunately, because of Teaching Practice requirements only thirty-two of these subjects were available for the tests of autonomy of imagery.

### III.—RESULTS.

Evaluation of the criterion movement was by three experienced gymnastic teachers from the teachers' college and the State university. The agreement between the ratings on a 0-10 scale in halves gave a reliability of .96, Ebel<sup>6</sup>. For each subject the mean of the ratings by the three judges was taken as the performance score which was then T Scaled to provide a criterion.

A factor analysis of the data from the Sutcliffe Battery indicated that the six modes examined in fact contained only one imagery factor with the possibility of a very minor second factor. This finding is in accord with those of

Sutcliffe.<sup>20</sup> However, as an earlier study Arnold<sup>1</sup> had found that a combination of visual and kinaesthetic imagery could predict body sway, the relation of these two modalities to the criterion was examined as well as the general imagery scores which were means of the scores obtained on each of the six modes. The correlations obtained are given in Table 1.

TABLE 1  
RELATION OF VIVIDNESS OF IMAGERY TO CRITERION SKILL SCORES.

Visual	Kinaesthetic	General
- .227	- .063	- .089

Not one of these first order correlations was significant at the 5 per cent. level. It would seem that vividness of imagery was not a critical factor in the efficiency of mental practice of physical skills. On the other hand, a multiple linear regression analysis which included these data with eleven other variables, suggested that kinaesthetic imagery was a minor factor in success at the criterion skill.\*

The sample for the Autonomy Testing had been reduced to thirty-two because of college commitments in the second week of testing. The remaining data were grouped in a high/low contingency table (Table 2).

TABLE 2  
RELATION BETWEEN AUTONOMY OF IMAGERY AND PERFORMANCE SCORES.

Imagery \ Performance	High (50.0+)	Low (49.9-)
Autonomous (18+) ....	5	8
Controlled (17-) .....	14	5

Using Chi Square, a value of 3.83 was obtained which was marginally short of the 3.84 required for significance at 5 per cent. level (one tailed). The Mann-Whitney U Test gave  $U=74$  which was significant at the 5 per cent. level (one tailed). Because of the reduced sample size for the Image Autonomy Test this factor was not included in the regression analysis. However, an attempt was made to look at the interaction of vividness and controllability of imagery and its relation to the performance scores. When the performance

\*In the regression equation of fourteen variables Kinaesthetic Imagery was the fifth variable giving  $R=.556$ . When allowances were made for intercorrelation of variables, Kinaesthetic imagery appeared as the second or third variable depending on whether Normality Coefficient, 'Students'  $t$ , or  $F$  level was taken as the criterion.



scores of subjects with vivid autonomous imagery were compared with those of subjects with all other imagery combinations, the data appeared as in Table 3.

TABLE 3

AN INTER-RELATION OF VIVIDNESS AND AUTONOMY OF IMAGERY AND ITS RELATION TO PERFORMANCE SCORES.

Performance Imagery	High (50.0+)	Low (49.9-)
Vivid-Autonomous . . . .	1	7
All other combinations . .	12	12

It would seem from this distribution that vivid imagery which is not under the control of the individual tends to be associated with low performance scores in this sample.

#### IV.—DISCUSSION.

It would seem that neither vividness nor controllability of imagery separately predicts performance scores at a physical skill which has been 'learned' by mental practice. Previous studies have noted the lack of relation between imagery and intellect (Galton<sup>9</sup>), comprehension (Thorndike<sup>21</sup>), discrimination, logical thinking, appreciation of literature (Betts<sup>2</sup>), mental multiplication (Thorndike<sup>21</sup>), children's abilities as seen by their teachers (Carey<sup>3</sup>), and school marks (Carey<sup>3</sup>).

Contrary to these findings is the viewpoint of Pear<sup>14</sup> who suggested that the physical prowess of such people as Pavlova (ballet), Fry (cricket) and Tilden (tennis) might be because of the efficiency of their imagery. Similar statements have been with regard to games playing (Whitely<sup>23</sup>) and memory tests (Vernon<sup>22</sup>); and Arnold<sup>1</sup> found that a composite of kinaesthetic and visual imagery predicted body sway in his experimental study. More particularly related to this study was Whiteley's<sup>23</sup> finding that improvement in ball throwing for accuracy after mental practice correlated .633 with scores on a test of imagery he used. Whitely used the ability to recall the contents of two trays as his imagery test and this memory recall might measure a difference trait from that governing the subject's ability to conjure up an image of something he has never seen before. Undermining this explanation is evidence that the difference between memory and imagery might not be significant as Vernon<sup>22</sup> stated that "The assumption is that the individual for whom one kind of imagery predominates will do best in the corresponding type of memory tests" (p. 195).

As a single variable, the controllability of imagery, appears to be more significant, but it is the interaction of these two facets of imagery which promises to be most informative. Certainly, subjects with uncontrollable vivid imagery appear to have low performance scores. This would support Clark's<sup>4</sup> observation that one of his subjects could not imagine bouncing a ball as the ball "would not bounce and stuck to the floor" (p. 567). This inability to manipulate his imagery disturbed the subject so much that he could not visualise the complete sequence of changing events. If such visualisation of the changing body positions during a gymnastic movement was necessary to benefit from mental practice, then it is easy to see why vivid autonomous imagery would be a handicap.

The controllability of vivid imagery might be part of a general ability to ignore irrelevant or extraneous stimuli, such as is required in the Gottschaldt Figures where one must break down a Pattern to locate within it a criterion memory pattern. As there appears to be a low significant correlation between Gottschaldt Figures B and scores on physical performance after mental practice (Start<sup>19</sup>), this ability could be some form of 'concentration' as has been suggested in personal correspondence with W. L. Steel (University of Manchester) discussing his work with ball throwing after mental practice.

An incidental point arising from this study is the result of the factor analysis of the data on image modes. It would appear that the hypothesis of audiles, visiles, etc., might have to be re-examined as a result of findings with the Sutcliffe Battery. Perhaps there are only good and bad imagers without any real variation in modality although from habit people might tend to use one imagery form more than another or favour one because of the ease with which it can be verbalised.

#### V.—REFERENCES.

- <sup>1</sup> ARNOLD, M. (1947). Brain function and emotion. *Amer. Psychologist*, 2, 344-345.
- <sup>2</sup> BETTS, C. H. (1909). *The Distribution and Functions of Mental Imagery*. Teachers' College, Columbia University Press.
- <sup>3</sup> CAREY, N. (1915). Factors in the mental processes of school children, 1—Visual and auditory imagery. *Brit. J. Psychol.*, 7, 453-490.
- <sup>4</sup> CLARK, L. V. (1960). Effect of mental practice on the development of a certain motor skill. *Research Quarterly*, 31, 560-569.
- <sup>5</sup> DAVIES, R. C. (1938). The relation of muscle action potentials to difficulty and frustration. *J. Exp. Psychol.*, 23, 141-158.
- <sup>6</sup> EBEL, R. L. (1951). Estimation of the reliability of ratings. *Psychometrika*, 16, 407-424.
- <sup>7</sup> FREEMAN, G. L. (1931). Studies in muscular tension. IV—Mental activity and the muscular process. *Psychol. Review*, 38, 428-449.
- <sup>8</sup> FREEMAN, G. L. (1931). The spread of muscular activity during mental work. *J. Gen. Psychol.*, 5-29, 479-484.
- <sup>9</sup> GALTON F. (1919). *Inquiries into Human Faculty and its Development*. Dent (Everyman's Edition), London (originally published in 1883).
- <sup>10</sup> GORDON, R. (1949). An investigation into the factors that favour the formation of stereotyped images. *Brit. J. Psychol.*, 39.3, 156-167.
- <sup>11</sup> HUMPHREY, G. (1948). *Directed Thinking*. New York: Dodd, Mead Co., p. 148.
- <sup>12</sup> JACOBSON, E. (1932). Electrophysiology of mental activities. *Am. J. Psychol.*, 49, 677-694.
- <sup>13</sup> MCKELLAR, P. (1957). *Imagining and Thinking*. London: Cohen and West.
- <sup>14</sup> PEAR, T. H. (1922). *Remembering and Forgetting*. London: Methuen.
- <sup>15</sup> PEAR, T. H. (1824). *Skill in Work and Play*. London: Methuen.
- <sup>16</sup> SHAW, W. A. (1940). The relation of muscle action potentials to imaginal weight lifting. *Arch. Psychol.*, 247, 3515-3545.
- <sup>17</sup> SMITHELLS, P. A., and CAMERON, P. E. (1962). *Principles of Evaluation in Physical Education*. New York: Harper.
- <sup>18</sup> START, K. B. (1960). The relationship between intelligence and the effect of mental practice on the performance of a motor skill. *Research Quarterly*, 31.4, 644-649.
- <sup>19</sup> START, K. B. (1963). *Special Perception and the Mental Practice of a Motor Skill* (an unpublished study. University of Western Australia).
- <sup>20</sup> SUTCLIFFE, J. P. (1962). *Progress Report on Project M.3950*. U.S. Public Health Service, Dept. Psychol., University of Sydney.
- <sup>21</sup> THORNDIKE, E. L. (1907). On the function of visual images. *J. Philos. Psychol. and Sci. Method*, 4, 12, 324-327.
- <sup>22</sup> VERNON, M. D. (1937). *Visual Perception*. Cambridge University Press.
- <sup>23</sup> WHITELEY, G. (1962). *The Effects of Mental Rehearsal upon a Motor Skill*. Dip. Ed. Thesis. Manchester University.
- <sup>24</sup> WILLIAMS, R. B. (1914). Analysis of forms of reaction movements. *Psychol. Monog.*, 17, 56-115.



# A COMPARISON OF PEER AND TEACHER PERSONALITY RATINGS OF PUPILS IN A SECONDARY MODERN SCHOOL

By H. J. HALLWORTH

(*Department of Education, University of Birmingham*)

AND A. MORRISON

(*Moray House College of Education, Edinburgh*)

**SUMMARY.** Previous work on teachers' personality ratings of secondary school pupils has established two main dimensions, one descriptive of the 'reliable and conscientious' or 'emotionally stable' pupil, the other of the 'extraverted' pupil. For the present study, ratings of pupils were obtained from both teachers and peers, using comparable lists of traits.

The subjects were 200 boys and girls in the first, second and third years of a co-educational secondary modern school. Ratings were made separately for boys and girls within each class, by the class teacher and by three pupils. The three pupils' ratings were combined to give one peer rating on each trait.

Using an electronic computer, a correlation matrix was obtained for each sex group within each class, the matrices for each sex were combined using *z*-scores, a principal components analysis was obtained for each combined matrix, and the axes were rotated to the Varimax criterion. In each case six factors were extracted, accounting for 68.9 per cent. of the variance in the boys' analysis and 60.3 per cent. in the girls' analysis.

Four of these factors are comparable across both analyses. They substantiate the earlier findings of two principal dimensions for the assessment of pupils by teachers, and indicate that peer assessments follow a similar pattern. A factor of non-academic leadership is discussed.

## I.—INTRODUCTION.

In several recent researches the well established techniques of personality rating have been used to investigate the assessments made by teachers of their pupils. The interest has been in the raters rather than the rated and the aim has been to study the ways in which teachers classify and label their pupils in order to impose form and meaning on the complex environment of the secondary school classroom (e.g., Hallworth, 1962, 1964).

Precautions have been taken to avoid a spurious halo effect. Teachers co-operating in these investigations have been well acquainted with their pupils; the groups of pupils have been of like sex and of approximately the same chronological age and intellectual ability; and all pupils have been rated on one trait before the teacher has proceeded to the next trait. For each trait the mean of the scale has been assumed to be the mean of the group being rated, and ratings have been distributed on a five-point scale in accordance with the percentages of the normal distribution. A variety of traits has been used in order to cover more than one dimension of assessment.

Two dimensions have been consistently obtained and have been named 'reliability and conscientiousness' or 'emotional stability'; and 'sociability' or 'social extraversion' (e.g., Hallworth, 1961). These may be dimensions of children's personality; they have been interpreted as dimensions by which teachers perceive and classify their pupils.

Little has been published, however, concerning the dimensions of pupils' personality ratings of their peers. Still less information is available regarding comparative studies of teacher and peer ratings of pupils. The present article describes an investigation in which both kinds of rating were used. The study was designed to extend the earlier work on teacher ratings by using the same traits, rating techniques and methods of analysis for both teacher and peer ratings.

## II.—DESIGN OF THE INVESTIGATION.

The subjects were 200 pupils of a co-educational secondary modern school in a working class suburb of a Midlands city. They were in two classes in each of the first, second and third years.

Each class teacher rated his pupils on twelve personality traits as shown in Table 1. These were the identical traits used in earlier studies, and were defined by the same brief descriptions. Ratings were also made of attainment in school work and ability in games. Boys and girls within each class were treated as separate groups and ratings were made on a five-point scale in the manner already described.

From each sex group within each class, three pupils were selected at random to make ratings of their peers on the twelve traits shown in Table 1. Eight of these traits were almost identical with eight on the teachers' list of traits. In the case of the remaining four traits on the teachers' list it was considered impracticable or undesirable for pupils to make ratings of their peers. Four other traits were, therefore, chosen which were judged to be of importance to children. The three sets of ratings obtained in this way for each peer group were combined to give a pooled rating on each trait.

To the two sets of ratings, seven other measures were added. These were the Children's Manifest Anxiety Scale (Casteneda, 1956), a measure of absenteeism, family size, ordinal position in family, sociometric status, a measure of attitude to school obtained by a Likert-type scale, and a measure of active and passive interests obtained by an interests test (Peel, 1948, 1952).

## III.—STATISTICAL ANALYSIS.

The analysis of the data was made with the aid of a Ferranti Mercury electronic computer. For each of the single-sex class groups a correlation matrix was computed, using the thirty-three measures described above. The six matrices for the boys' groups were combined, using *z*-scores, to obtain a matrix of average correlations. This procedure was repeated for the matrices from the girls' groups. A principal components analysis was then obtained for each of the two matrices of average correlations. Six components were extracted in each case, accounting for 68.93 per cent. of the variance in the boys' analysis, and for 60.34 per cent. in the girls'. Finally, the axes were rotated to the Varimax criterion, Kaiser's loose limit being used (Kaiser, 1959).\*

## IV.—RESULTS.

The first four Varimax factors in both analyses were very similar. They are shown in Table 2, together with the percentage of variance extracted by each.

\* The computer programs were written by H. J. Hallworth in mixed Mercury Autocode and machine code.



(a) *Factor I: Peer ratings: Extraversion.*

Varimax Factor I accounts for 21.05 per cent. of the variance in the boys' analysis and 14.07 per cent. in the girls'. It has high loadings as follows:

	Boys	Girls		Boys	Girls
15. Friendliness .....	.82	.80	22. Grown-up attitudes ..	.54	.43
16. Leadership .....	.64	.59	23. Popularity .....	.88	.74
17. Persistence .....	.54	.48	24. Confidence .....	.81	.58
19. Cheerfulness .....	.82	.77	25. Attitude to School ..	.54	.31
20. Sociability .....	.84	.85	26. Active interests ....	.71	.45
21. Aggressiveness .....	.52	—	31. Sociometric score ...	.74	.49

All peer ratings have high loadings except 18. Co-operation, with teachers and, in the girls' analysis, 21. Aggressiveness. Sociometric status also has a high loading. The factor obviously represents peer evaluation, and shows a halo effect which is greater among the boys than among the girls. This finding is consistent with evidence that social groupings among boys tend to be larger and more hierarchical than among girls.

The heaviest loadings, however, are on the ratings of 20. Sociability, 23. Popularity, 15. Friendliness, and 19. Cheerfulness. It is, therefore, justifiable to regard the factor as in some ways representing, among peer ratings, the equivalent of the sociability or social extraversion factor found in earlier studies of teacher ratings.

The only teacher ratings with loadings worthy of notice are in the boys' analysis, viz., 10. Popularity (.46) and 14. Games ability (.36).

(b) *Factor II: The good pupil.*

The second Varimax factor accounts for 17.08 per cent. of the variance in the boys' analysis and 12.71 per cent. in the girls'. Highest loadings are as follows:

	Boys	Girls		Boys	Girls
1. Emotional Stability ..	.73	.54	17. Persistence .....	.64	.60
2. Trustworthiness ....	.82	.81	18. Co-operation .....	.73	.61
3. Persistence .....	.75	.62	22. Grown-up attitudes ..	.47	.42
4. Co-operation .....	.74	.63	25. Attitude to school ..	.66	.59
			28. Absenteeism .....	-.36	-.53

The first four of these variables are teacher ratings of precisely those traits previously found to lie on the dimension of 'reliability and conscientiousness' or 'emotional stability,' viz., 1. Emotional stability, 2. Trustworthiness, 3. Persistence, and 4. Co-operation with teachers. The next four are peer ratings, two of them being identical with two of the teacher ratings, viz., 17. Persistence and 18. Co-operation with teachers. The factor obviously represents those characteristics which are assumed to be the marks of a good pupil.

The measure of absenteeism was based upon the number of periods of absence rather than total absences, and was intended to give an indication of casual absenteeism. It had previously been found to give an indirect assessment of poor morale in both school (Hallworth, 1953) and industry (Mayo, 1945). It is, therefore, significant that the measure has negative loadings in both analyses. The boy or girl who was perceived by both teachers and peers as a good pupil tended to have a low rate of absenteeism.

In the boys' analysis there are significant loadings on the teacher ratings of 9. Maturity (.65) and 13. Attainment in school work (.60), and on the peer rating of 16. Leadership (.41).

(c) *Factor III: Teacher Ratings: Extraversion.*

Varimax Factor III accounts for 13.16 per cent. of the variance in the boys' analysis and 14.35 per cent. in the girls' analysis. Highest loadings are as follows:

	Boys	Girls		Boys	Girls
5. Cheerfulness .....	.76	.75	10. Popularity .....	.46	.69
6. Humour .....	.84	.77	11. Confidence .....	.60	.28
7. Sociability .....	.76	.86	14. Games .....	.23	.33
8. Self-assertion .....	.66	.69	19. Cheerfulness .....	.21	.24
12. Spontaneity .....	.86	.80	20. Sociability .....	.23	.18
			26. Active interests ....	.20	.27

The first five of the variables with high loadings are the teacher ratings on those traits which in earlier studies have defined the dimension of extraversion, viz., 5. Cheerfulness, 6. Sense of humour, 7. Sociability, 8. Self-assertion, and 12. Spontaneity.

The two teacher ratings which have previously been found to lie between the dimension of 'reliability and conscientiousness' and the dimension of 'extraversion,' viz., 10. Popularity, and 11. Confidence, have moderate loadings on Factor III, as would be expected. From an examination of the loadings on Factor II, it is apparent that, in general, these variables again lie between the two dimensions from the teacher ratings. The teacher rating of 14. Ability in Games has low positive loadings on Factor III in both analyses.

Three peer ratings have low positive loadings in both analyses, viz., 19. Cheerfulness, 20. Sociability, and 26. Active Interests; whilst 21. Aggressiveness, and 16. Leadership have moderate loadings in the girls' analysis. There is, in effect, an indication of some slight overlap between teacher and peer perceptions of the sociable or extraverted boy and girl.

(d) *Factor IV: Non-Academic Leadership.*

Varimax Factor IV accounts for 6.31 per cent. of the variance in the boys' analysis and 7.07 per cent. in the girls'. Highest loadings are:

	Boys	Girls		Boys	Girls
9. Maturity .....	.50	.70	16. Leadership .....	.30	.42
11. Confidence .....	.40	.33	22. Grown-up attitudes .	.35	.36
14. Games .....	.53	.67	24. Confidence .....	.30	.26
			26. Active interests ....	.35	.43



The first three variables are teacher ratings, the last four are peer ratings. The loadings are relatively low, but the consistency for both analyses and for similar traits rated by teachers and peers suggests that they represent a dimension not previously identified. Tentatively, it may be named 'non-academic leadership.' The picture is of a socially mature, confident boy or girl who is physically active and good at games, and who is regarded by his peers as a leader. His teachers tend to assess him as 1. Emotionally stable ( $\cdot 22$ ,  $\cdot 23$ ), but lacking in 4. Co-operation with teachers ( $-\cdot 24$ ,  $-\cdot 25$ ).

A boy who is regarded as a non-academic leader also tends to be assessed by his peers as lacking in 18. Co-operation with teachers ( $-\cdot 30$ ); and by both teachers and peers as 8. Self-assertive ( $\cdot 45$ ) and 21. Aggressive ( $\cdot 53$ ).

#### V.—CONCLUSION.

The two outstanding features of the results of this investigation are the consistency of the factor structure for both boys' and girls' analyses, and the consistency with earlier findings insofar as the teacher ratings are concerned. Once again it was found that teachers make personality ratings of their pupils along two dimensions which may be termed 'conscientiousness and reliability' and 'extraversion.'

The peer ratings were also made chiefly along two main dimensions. The first of these, Factor I, showed a general halo effect which was more pronounced among the boys than among the girls: the only peer rating with low loadings on both analyses was 18. Co-operation with teachers, although 21. Aggressiveness had a low loading in the girls' analysis. Sociometric score had a high loading on both analyses, but all teacher ratings and all other variables had low loadings. The highest loadings on this dimension were on the peer ratings of 'extraversion' traits. It was, therefore, assumed to be the equivalent of the teachers' dimension of 'extraversion.' The greater halo effect among peer ratings may be due to the relative lack of verbal sophistication of the younger raters, or to a greater subjective element in judgments since they are part of the social group whose members are being rated.

The second main dimension for the peer ratings was Factor II, and heavy loadings were obtained on 16. Persistence, 17. Co-operation with teachers, 22. Grown-up attitudes, and 25. Positive attitudes to school. This was the factor with heavy loadings for the teacher ratings of 'reliability and conscientiousness.' In the boys' analysis, there were high loadings on the teacher ratings of 9. Maturity, and 13. Attainment in school work, and on the peer rating of 16. Leadership. Teachers and children agreed upon the personality traits of a 'good pupil.' The measure of 28. Absenteeism had sizeable negative loadings in both analyses, as would be predicted.

Factor III represented the teacher rating dimension of 'extraversion.' This factor had low but consistent positive loadings on the peer ratings of 19. Cheerfulness, 20. Sociability and 26. Active interests, indicating a small amount of overlap in teacher and peer perceptions.

A new dimension of 'non-academic leadership' was indicated by Factor IV. Both teachers and peers agreed that a pupil of this kind is mature or grown-up in attitude, and confident; he is also good at games and has active interests and, if a boy, he is self-assertive and aggressive, and has a low score on the Scale of Manifest Anxiety.

The consistency of these results suggests that they represent regularities in teacher and peer assessments of personality which have not previously been demonstrated. The apparent differences between teacher and peer perceptions invite further investigation.

TABLE 1

Traits Rated by Teachers	Traits Rated by Peers
Emotional Stability.	Friendliness
Trustworthiness.	Leadership.
Persistence.	Persistence.
Co-operation with Teachers.	Co-operation with Teachers.
Cheerfulness.	Cheerfulness.
Sense of Humour.	Sociability.
Sociability.	Aggressiveness.
Self-assertion.	Grown-up Attitudes.
Maturity.	Popularity.
Popularity.	Confidence.
Confidence.	Positive Attitude to School.
Spontaneity.	Active Interests.

TABLE 2

LOADINGS OF VARIMAX FACTORS OBTAINED FROM COMBINED CORRELATION MATRICES.

	Factor I		Factor II		Factor III		Factor IV	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
1. Emotional stability ..	.06	-.07	.73	.54	.11	-.31	.22	.23
2. Trustworthiness .....	.08	-.07	.82	.81	.25	-.02	-.02	-.03
3. Persistence .....	.10	-.12	.75	.62	.21	-.01	.12	.51
4. Co-operation .....	.00	-.21	.74	.63	.39	.34	-.24	-.25
5. Cheerfulness .....	.24	.06	.33	.14	.76	.75	-.01	-.06
6. Humour .....	-.01	.25	.11	.00	.84	.77	-.06	-.01
7. Sociability .....	.10	.09	.33	.06	.76	.86	-.07	.00
8. Self-assertion .....	.17	.17	-.05	-.27	.66	.69	.45	.19
9. Maturity .....	.26	-.12	.65	.27	.04	-.16	.50	.70
10. Popularity .....	.46	.21	.30	.25	.46	.69	.10	.19
11. Confidence .....	.19	.16	.26	.04	.60	.28	.40	.33
12. Spontaneity .....	.18	-.04	.08	-.08	.86	.80	.10	.05
13. Attainment .....	.08	.14	.60	.35	.19	.20	.29	.45
14. Games .....	.36	.16	.19	-.15	.23	.33	.53	.67
15. Friendliness .....	.82	.80	.16	.14	.10	-.06	-.15	-.10
16. Leadership .....	.64	.59	.41	-.05	.17	.35	.30	.42
17. Persistence .....	.54	.48	.64	.60	.14	-.10	-.03	.11
18. Co-operation .....	.39	.21	.73	.61	.01	-.24	-.30	.12
19. Cheerfulness .....	.82	.77	-.01	-.15	.21	.24	-.03	.03
20. Sociability .....	.84	.85	.06	.09	.23	.18	.04	-.01
21. Aggressiveness .....	.52	.26	-.03	-.29	.08	.51	.53	.09
22. Grown-up attitudes ..	.58	.43	.47	.42	-.05	-.21	.35	.36
23. Popularity .....	.88	.74	.06	.32	.02	.20	.22	-.05
24. Confidence .....	.81	.58	.14	-.02	.05	.22	.30	.26
25. Attitude to school ..	.54	.31	.66	.59	-.04	.01	-.04	-.01
26. Active Interests .....	.71	.45	.14	-.09	.20	.27	.35	.43
27. C.M.A.S. ....	-.16	.10	.23	-.21	-.22	-.05	-.24	-.03
28. Absenteeism .....	.03	-.08	-.36	-.53	.01	-.05	.00	-.08
29. Family size .....	.04	.01	-.13	-.19	-.02	-.08	.02	.00
30. Ordinal position .....	.06	.20	-.03	-.05	.10	.01	-.03	.09
31. Sociometric score .....	.74	.49	.21	.40	.03	.08	-.01	-.07
32. Attitude to school ..	.06	-.19	-.13	.21	.17	.06	-.15	.11
33. Active interests test ..	-.02	.02	.12	.05	-.28	.09	.10	.04
Percentage of Total Variance .....	21.05	14.07	17.08	12.71	13.16	14.35	6.31	7.07

Variables 1-14, inclusive, are teachers' ratings.  
 Variables 15-26, inclusive, are peer ratings.



## VI.—REFERENCES.

- CASTENEDA, A., McCANDLESS, B. R., and PALERMO, D. S. (1956). The children's form of the manifest anxiety scale. *Child Dev.*, 27, 317-326.
- HALLWORTH, H. J. (1953). Sociometric relationships among grammar school boys and girls. *Sociometry*, 16, 39-70.
- HALLWORTH, H. J. (1961). Teachers' personality ratings of high school pupils. *J. Educ. Psychol.*, 52, 297-302.
- HALLWORTH, H. J. (1962). A teacher's perception of his pupils. *Educ. Rev.*, 14, 124-133.
- HALLWORTH, H. J. (1964). Personality ratings of adolescents: a study in a comprehensive school. *Brit. J. Educ. Psychol.*, 34, 171-177.
- KAISER, H. F. (1959). Computer program for Varimax rotation in factor analysis. *Educ. Psychol. Meas.*, 19, 413-420.
- MAYO, E. (1945). *The Social Problems of an Industrial Civilisation*. Boston: Harvard Univ. Press.
- PEEL, E. A. (1948). Assessment of interest in practical topics. *Brit. J. Educ. Psychol.*, 18, 41-47.
- PEEL, E. A. (1952). Assessment of children's interests. *Educ. Rev.*, 4, 164-169.

(Manuscript received 9th April, 1964).

# THE INCIDENCE AND NATURE OF MALADJUSTMENT AMONG CHILDREN IN SCHOOLS FOR THE EDUCATIONALLY SUBNORMAL

BY MAURICE CHAZAN

(*Department of Education, University College of Swansea*)

**SUMMARY.** An investigation has been carried out into the incidence, nature and aetiology of maladjustment among children in special schools for E.S.N. children in South Wales. This article discusses the findings of the first stage of the enquiry, in which detailed information was obtained about the incidence and nature of maladjustment, in the school situation, among 169 E.S.N. pupils and the same number of children in ordinary schools, matched for age, sex and socio-economic background. The E.S.N. pupils were a complete sample of two age groups (9+ to 10+ and 13+ to 14+) in eight special schools.

The E.S.N. pupils came predominantly from the lower social strata. They had more physical ailments and defects than the controls. On the basis of the Bristol Social Adjustment Guide, over a third of the E.S.N. children were maladjusted, nearly three times as many as in the control group. The E.S.N. children showed considerably more symptoms of depression, hostility towards adults, inhibition and emotional tension than the controls. There was also a higher incidence of speech defects, unsatisfactory attendance at school, delinquency and behaviour associated with delinquency in the E.S.N. sample. Maladjustment among the E.S.N. children was often associated with low sociometric status.

These findings underline the importance of attention to the emotional and social needs of E.S.N. children.

## I.—INTRODUCTION.

THE need for the recognition of the multiple nature of the handicap of educationally subnormal children is now being emphasized (Ministry of Education, 1962). Several writers have commented on the tendency for intellectual subnormality to be accompanied by emotional immaturity or maladjustment (Burt, 1937; Walker, 1950; McLachlan, 1955). Tansley (1951) stresses that an educational programme for E.S.N. children which does not meet emotional needs is bound to fail and suggests that the incidence of emotional disturbance in E.S.N. children is considerably higher than that found among children in general. We lack, however, detailed knowledge of the emotional and social problems of children in schools for the E.S.N., and without this knowledge there is uncertainty about the extent of the problem of maladjustment amongst E.S.N. pupils, the methods of special educational treatment which are most suitable, and the kinds of ancillary services which would help the schools.

To attempt to throw some light on these questions, an enquiry was planned into the incidence, nature and aetiology of maladjustment in special schools for E.S.N. children in South Wales. This article describes the investigation and presents the findings of the first stage of the enquiry.

## II.—PREVIOUS WORK.

There have been few systematic studies of the extent and nature of maladjustment among E.S.N. children. Burt (1937) found that about one-third of the backward children he studied in the London area were temperamentally handicapped. Tansley and Gulliford (1960) point out that E.S.N. children are more suggestible and more easily led astray than normal children; their means of finding emotional satisfaction and expression are less varied, and their emotional development is slower and more uneven. Stott (1961) emphasizes



that backward children often suffer from a defect of motivation, and in particular are timid and unassertive. Studying 153 backward children in the Bristol region, he found 45 per cent. to be unforthcoming, 20 per cent. unsettled, and 17 per cent. with other behaviour disturbance; only 24 per cent. were rated as stable.

Several American writers (Walker, 1950; McLachlan, 1955; Hutt and Gibby, 1958) stress that the subnormal child experiences considerable emotional difficulties in his early years. Feelings of inadequacy may result from delay in reaching developmental milestones, with consequent parental anxiety. Independence from parents may be achieved at a later stage than is normal; the subnormal child cannot adequately or realistically assess the demands of the external world; he finds it more difficult to control his impulses and to learn to substitute future gratifications for immediate pleasures; and he tends to be insecure and to have excessive fears and anxieties. The mentally retarded child is not only more prone to maladjusted behaviour reactions than normal children, but such behaviour tends to persist over a much longer period of time.

The relationship between maladjustment and school failure has been discussed by a number of writers (Duncan, 1947; Pringle and Bossio, 1958; Chazan, 1962). The cumulative effects of the lack of success which characterizes the E.S.N. child's efforts in his early years at school tend to produce unsatisfactory attitudes towards the learning process, a low level of aspiration, and symptoms of emotional disturbance at home or in school.

Feelings of inferiority and failure with regard to school work may be exacerbated by the inability of the backward child to obtain acceptance by his peers in the ordinary school. Sociometric studies in the U.S.A. (Johnson, 1950; Johnson and Kirk, 1950; Miller, 1956; Dunn, 1963) indicate that, while the educable mentally retarded in the ordinary school obtain lower social status positions than their intellectually superior peers, they are not rejected by their classmates because of their slowness in learning but because of unacceptable behaviour patterns.

### III.—AIMS AND METHODS OF THE INVESTIGATION.

The aims of the enquiry were (i) to ascertain the incidence and nature of maladjustment in the school situation in a representative sample of children attending schools for the E.S.N. in South Wales, and to compare the findings with those from a matched control group of children attending ordinary schools (Stage I); (ii) to discover, by a close study of a smaller number of pupils, what factors are associated with maladjustment in E.S.N. children (Stage II).

The first stage of the enquiry was begun in May, 1962, when the Bristol Social Adjustment Guide, *The Child in School* (Stott, 1963), was completed by the teachers of 169 E.S.N. pupils (110 boys, 59 girls) for each child. Bristol Guides were also completed for the same number of children in ordinary schools, matched with the E.S.N. pupils for age, sex and socio-economic background. As it is important to control the age factor in studies of maladjustment, and as numbers too large for the resources of this enquiry would have been involved if children from each age group at school had been included, it was decided to restrict the investigation to two age groups, 9+ to 10+ and 13+ to 14+, the usual age ranges of children in the third years of the ordinary junior and secondary school. This made it possible to include these two complete age groups in eight special schools in South Wales (six day, two residential); the schools had a total school population of approximately 800 pupils. The control group was selected from 496 pupils of the same age ranges attending ordinary junior and secondary schools in South Wales. These pupils constituted the sample in a



simultaneous enquiry into the incidence and nature of maladjustment amongst children in ordinary schools (Chazan, 1963, a).

In addition to the information from the Bristol Guide, data were obtained about each E.S.N. pupil's intelligence, school progress, home background and sociometric status by means of a supplementary questionnaire completed by each head teacher, reference to Form 2 H.P., which gave the results of one or more individual intelligence tests, and a sociometric test. As children in special schools for the E.S.N. are regularly given individual intelligence and attainment tests, the writer wished to avoid, as far as possible, giving additional tests of this kind; there were, therefore, some gaps in the information available concerning basic attainment levels.

The difficulties of research in this field are great. There is a lack of reliable objective instruments for the assessment of maladjustment, which in any case is essentially a subjective concept (Chazan, 1963, b). Individual interviews are necessary when information is required from the pupils themselves, as written personality tests and questionnaires cannot be used with children who have a low level of attainment in reading and limited verbal comprehension. The Bristol Social Adjustment Guide seemed, for a number of reasons, to be the most appropriate and satisfactory instrument for obtaining an overall picture of the incidence and nature of maladjustment as well as detailed information about individual children. Firstly, the main aim of the enquiry was to find out how the teachers saw their pupils and what barriers the children presented to them. Secondly, although the subjective views of the individual teacher must of necessity affect the assessment, teachers do complete the guides consistently and skilfully, especially when they know the child well, and, as Burt (1937) pointed out, the psychologist must rely a good deal on the daily observation of the trained and experienced teacher in the classroom. Thirdly, although the Guide needs further statistical analysis (Vernon, 1964), its reliability and validity have been established by a number of studies making extensive use of it (Stott, 1963). Petrie (1962), for example, found the reliability and validity of the Bristol Guides in the assessment of maladjustment to be better than that of projection tests, and concluded that they provide a reasonably good means of assessing the effects of treatment.

#### IV.—DESCRIPTION OF SAMPLE.

(a) *Level of Intelligence.* I.Qs. based on individual intelligence tests, mainly the Terman-Merrill Scale, were available for all except two of the E.S.N. pupils. In most cases, the results of two or more tests carried out at different times had been recorded, and as these results showed the fluctuations common among E.S.N. children (Stott, 1960), the mean of the I.Qs. was calculated for each child. The distribution of I.Qs. does not differ greatly from that found in schools for E.S.N. pupils in predominantly rural regions (Ministry of Education, 1962). 17.8 per cent. of the E.S.N. sample had I.Qs. below 60, 42.6 per cent. had I.Qs. between 60 and 69, and 38.4 per cent. had I.Qs. of 70 or over. It is worthy of note that the population of special schools for the E.S.N. can no longer be thought of in terms of the conventional 50-70 I.Q. range. There was a higher proportion of pupils with I.Qs. over 70 in the younger age range studied; 31 per cent. of the younger girls and 54 per cent. of the younger boys fell in this category, as compared with 17 per cent. and 42 per cent. of the older girls and boys respectively. This may be an artefact of the tests used; Rushton and Stockwin (1963) found that there was a statistically significant decline in the I.Qs., as measured by the Terman-Merrill Scale (Form L) of E.S.N. boys between the ages of 9 and 15 years.



(b) *Level of Attainment.* The information provided by the schools on the basic attainments of the E.S.N. children showed the formidable task which the teacher in the special school faces in trying to improve the scholastic attainments of E.S.N. pupils.

Few of the younger children had reading or arithmetic ages above 7 years, and even in the older age group, over one-half of the boys and nearly 45 per cent. of the girls were 'illiterate' (with a reading age below 7), while a third of both the senior boys and girls had arithmetic ages below 7. The low ceiling of attainment reached by most E.S.N. pupils even after many years of special education has been shown by Moran (1960), who found the mean attainment ages of 300 subnormal adolescents (aged 14-16+) in special schools in the London area to be: reading age 8 years 3 months; arithmetic age, 7 years 6 months.

(c) *Physical Condition and Defects.* The sections on general health and physical defects in the Bristol Guide indicate the problems presented by the pupil to the teacher because of poor health or physical defect. Table 1 shows that compared with the control group significantly fewer of the E.S.N. children as a whole enjoyed good health and freedom from physical defect, defined by the absence of any adverse items underlined in these sections of the Guide.

TABLE 1  
GENERAL HEALTH AND PHYSICAL DEFECTS IN E.S.N. AND CONTROL SAMPLES.

Group		Total N	In good general health		Free from physical defect	
			%	$\chi^2(\text{d.f.1})$	%	$\chi^2(\text{d.f.1})$
Boys						
Junior ..	E.S.N.....	50	56.0	14.17*	78.0	3.50
	Controls ..	50	94.0		94.0	
Senior....	E.S.N. ..	60	68.3	6.86*	76.7	0.76
	Controls ..	60	90.0		85.0	
Total ....	E.S.N. ....	110	62.8	21.33*	77.3	1.58
	Controls	110	91.8		89.1	
Girls						
Junior ..	E.S.N.....	23	56.5	4.00†	65.2	1.78
	Controls	23	87.0		91.3	
Senior....	E.S.N.....	36	22.2	15.65*	72.2	0.70
	Controls	36	69.4		77.8	
Total ....	E.S.N. ....	59	35.8	22.15*	69.5	2.67
	Controls ..	59	76.3		83.1	
Total .... (B. and G.)	E.S.N.....	169	54.4	44.80*	75.3	7.74*
	Controls ..	169	87.6		87.0	

\* Significant at 1 per cent. level or beyond.

† Significant at 5 per cent. level.

The most commonly occurring ailments among the E.S.N. children were frequent colds, infected ears, skin troubles and headaches; the most frequently reported defects were bad eyesight, poor hearing and bad co-ordination. The junior E.S.N. boys and the senior E.S.N. girls were particularly prone to ailments of various kinds.

(d) *Socio-economic background.* Only 7.2 per cent. of the children in special schools for the E.S.N. came from homes where the father was in an occupation in Social Classes I or II or in a white-collared occupation in Class III (Registrar-General's Classification of Occupations, 1960). 45.6 per cent. of the sample came from homes where the father was an unskilled labourer or unemployed, or where there was no father because of separation, desertion or death. The fathers of 46.7 per cent. of the children were skilled or semi-skilled manual workers (Classes III and IV). Stein and Susser (1960, b), in their sample of young adults who had been ascertained as E.S.N. at school, also found that the fathers were mainly semi-skilled, unskilled or casual workers (83 out of 102 fell into these categories). It may be that middle-class parents are more resistant than working-class parents to the idea of special schooling, and attempt to make their own arrangements for helping their children if they are making little educational progress.

Nearly one-half of the E.S.N. sample, which contained very few only children, came from families with five or more children; almost 10 per cent. of the children were from families with nine or more children. As Douglas and Blomfield (1958) have pointed out, children in the larger families, which tend to be found among the lower social classes, are often at a disadvantage in spite of increased welfare provision. Thus, E.S.N. children coming from poor homes and large families may be over-represented in special schools because they need residential care or present behaviour problems which makes their removal from an ordinary school a matter of urgency.

#### V.—INCIDENCE OF MALADJUSTMENT.

Table 2 shows the incidence of maladjustment in the school situation among the E.S.N. children on the basis of the Bristol Social Adjustment Guide, compared with that found in the control group of children in ordinary schools. As suggested by Stott (1963), each adverse item marked on the Guide was scored as one point. Children scoring less than ten points were rated 'stable', those scoring between ten and nineteen points were considered as 'unsettled', and those with twenty or more points were rated 'maladjusted.' The actual score is of less importance than the qualitative assessment of the child, but the scoring scheme is useful for the purpose of comparison between groups. For the purposes of statistical comparisons using  $\chi^2$ , these three categories were reduced to two: 'maladjusted' (20+ points) and 'not maladjusted' (0-19).

On this basis, over a third of the E.S.N. children were maladjusted, and less than a third fell into the 'stable' category. Both the E.S.N. boys and girls were significantly more maladjusted than the controls, and in only one of the sub-groups (Junior Girls, where the numbers were small) the difference between the E.S.N. and control groups with regard to the proportion of maladjusted children did not reach the 5 per cent. or 1 per cent. level of significance.

Child guidance clinics as well as a number of research studies (Stott, 1963) commonly find that a higher proportion of boys are maladjusted than girls.



This is the case in the present enquiry as far as the control group is concerned, though the difference is not statistically significant. The E.S.N. children, however, do not show any such sex difference in the incidence of maladjustment, and in fact, the most maladjusted group were the Senior E.S.N. girls, of whom 50 per cent. were categorized as maladjusted.

There were no significant differences between the Junior and Senior groups in the case of either the controls or E.S.N. children, and no over-all significant difference between schools, though the incidence of maladjustment ranged from 17 per cent. to 60 per cent. in the eight schools. The greatest problem of maladjustment was found in a residential school for senior E.S.N. boys, a day school for senior E.S.N. girls, and an all-age school for both boys and girls.

TABLE 2

INCIDENCE OF MALADJUSTMENT IN E.S.N. AND CONTROL GROUPS.

Group		N	Rating on Bristol Social Adjustment Guide			$\chi^2$ (d.f.1)	Mean Score
			Stable (0-9) %	Unsettled (10-19) %	Maladjusted (20+) %		
Boys							
Junior	E.S.N.	50	26.0	40.0	34.0	4.50†	16.7
	Controls	50	56.0	30.0	14.0		9.7
Senior	E.S.N.	60	31.7	36.7	31.7	4.55†	16.5
	Controls	60	53.3	31.7	15.0		9.6
Total	E.S.N.	110	29.1	38.2	32.7	10.00*	16.6
	Controls	110	54.5	30.9	14.5		9.6
Girls							
Junior	E.S.N.	23	47.8	21.7	30.4	3.12	13.5
	Controls	23	74.0	21.7	4.3		8.1
Senior	E.S.N.	36	25.0	25.0	50.0	9.39*	19.1
	Controls	36	61.1	27.8	11.1		8.3
Total	E.S.N.	59	33.9	23.7	42.4	15.38*	17.6
	Controls	59	66.1	25.4	8.5		8.2
Total (B. and G.)	E.S.N.	169	30.8	33.1	36.1	24.24*	16.7
	Controls	169	58.6	29.0	12.4		9.2

\* Significant at 1 per cent. level or beyond.

† Significant at 5 per cent. level.

## VI.—NATURE AND SYMPTOMS OF MALADJUSTMENT.

(a) *Patterns of maladjustment.* The contribution of the different syndromes assessed on the Bristol Guides was as follows, taking into account all the 'adverse' items marked in respect of the E.S.N. children :

	Percentage of Total
1. Depression (D) .....	16.4
2. Hostility towards adults (HA) .....	13.8
3. Unforthcomingness (U) .....	13.4
4. Anxiety about adult interest and affection (XA) .....	10.8
5. Unconcern for adult approval (K) .....	10.8
6. Restlessness (R) .....	9.5
7. Symptoms of emotional tension (M) .....	8.2
8. Anxiety for approval of and acceptance by other children (XC) ....	6.7
9. Withdrawal (W) .....	4.5
10. Hostility towards children (HC) .....	3.9
11. Miscellaneous nervous symptoms (MN) .....	2.6

To compare the types of disturbed behaviour found in the various groups of the investigation, the underlining of four or more adverse items was taken as suggesting a significant degree of disturbance in a particular syndrome, as suggested by Stott (private communication). On this basis, the E.S.N. children showed significantly more depression, hostility towards adults and inhibition than the controls, as well as more symptoms of emotional tension. Depression was shown particularly by the senior E.S.N. girls and boys, and inhibition by the junior boys and senior girls. The junior E.S.N. boys and girls also showed more restlessness and anxiety than the controls. There was little hostility towards other children and extremely withdrawn behaviour was rare.

Of the sixty-one maladjusted children in the E.S.N. sample, thirty-eight showed a predominantly aggressive pattern of behaviour, and sixteen withdrawn behaviour ; the other seven children displayed a mixture of both types of behaviour.

(b) *Unsatisfactory attendance at school.* Frequent or long absences from school may indicate or accompany maladjustment, or may result in the pupil finding it difficult to settle down in school when he is present. Although actual truancy appears to have been rare, a not inconsiderable proportion of the E.S.N. pupils, especially the senior girls, did not attend school regularly (Table 3). Teachers frequently commented that a pupil would have made much better progress if his attendance had been more regular.

(c) *Delinquency and Delinquency Proneness.* Table 3 also shows the number of pupils with a record of delinquency in or around school. The highest percentage of delinquents was found amongst the Senior E.S.N. boys. Significantly more of the E.S.N. pupils as a whole were delinquents than the controls. The main type of delinquency committed was stealing.



TABLE 3

INCIDENCE OF UNSATISFACTORY ATTENDANCE AT SCHOOL, TRUANCY AND DELINQUENCY  
IN E.S.N. AND CONTROL SAMPLES.

Group		Total N	Percentage of pupils with :					
			Unsatisfactory attendance at school		Record of truancy		Record of delinquency in or around school	
Boys			%	$\chi^2$ (d.f.1)	%	$\chi^2$ (d.f.1)	%	$\chi^2$ (d.f.1)
Junior	E.S.N.	50	36.0	n.s.	2.0	n.s.	8.0	n.s.
	Controls	50	22.0		0.0		4.0	
Senior	E.S.N.	60	31.7	n.s.	6.7	n.s.	20.0	n.s.
	Controls	60	21.7		6.7		8.3	
Total	E.S.N.	110	33.6	4.33†	4.6	n.s.	14.5	3.86†
	Controls	110	21.8		3.6		6.4	
Girls								
Junior	E.S.N.	23	13.1	n.s.	0.0	n.s.	8.7	n.s.
	Controls	23	17.4		0.0		0.0	
Senior	E.S.N.	36	38.9	n.s.	8.3	n.s.	4.4	n.s.
	Controls	36	25.0		0.0		0.0	
Total	E.S.N.	59	28.8	n.s.	5.1	n.s.	5.1	n.s.
	Controls	59	22.0		0.0		0.0	
Total (B. and G.)	E.S.N.	169	32.0	4.43†	4.7	n.s.	11.2	6.00*
	Controls	169	21.9		2.4		4.1	

\* Significant at 1 per cent. level or beyond.

† Significant at 5 per cent. level or beyond.

TABLE 4

DELINQUENCY PRONENESS SCORES OF E.S.N. AND CONTROL SAMPLES.

Group		Total N	Mean delinquency proneness score (Stott, 1961)	Significance of difference
Boys				
Junior ....	E.S.N. ....	50	10.23	5% level
	Controls ....	50	3.1	
Senior ....	E.S.N. ....	60	13.12	1% level
	Controls ....	60	6.52	
Total ....	E.S.N. ....	110	11.81	1% level
	Controls ....	110	4.97	
Girls				
Junior ....	E.S.N. ....	23	8.83	n.s.
	Controls ....	23	4.39	
Senior ....	E.S.N. ....	36	11.33	5% level
	Controls ....	36	5.5	
Total ....	E.S.N. ....	59	10.35	1% level
	Controls ....	59	5.07	
Total (Boys and Girls)	E.S.N. ....	169	10.71	.1% level
	Controls ....	169	5.01	

Use of the Delinquency Prediction Instrument (Stott, 1961), which is based on the Bristol Social Adjustment Guides and is still in the early stages of validation, showed a much greater tendency on the part of E.S.N. pupils to exhibit behaviour which is characteristically associated with delinquency (Table 4). The level of significance of the differences between means was calculated by non-parametric methods, using the sign test (Ferguson, 1959).

(d) *Speech Defects.* A high proportion of the E.S.N. children, particularly the boys, suffered from speech defects, which are not infrequently a symptom or accompaniment of emotional disturbance (Table 5). The main types of defect were stuttering (twenty-four children), mumbling, inaudible speech (twenty-four children) and babyish articulation (twelve children).

(e) *Social Isolation.* The sociometric status of each E.S.N. pupil was assessed by giving a sociometric test to each pupil of every class of which the children in the sample were members. The children were asked to give three choices for each of three criteria (seating, working, and playing companions), the teacher giving the test individually and orally where necessary. The rank correlations between maladjustment score on the Bristol Guide and sociometric status were calculated for each group of E.S.N. pupils by classes. These ranged



from  $-0.26$  to  $+0.97$ , but most of the correlations were positive (moderate or high), suggesting, as found by Belfield (1963), that the most accepted children tended to have good social adjustment and the least accepted children to have poor social adjustment.

Of the sixty-one maladjusted E.S.N. pupils, forty-six were 'below average' in sociometric status (0-8 choices), including thirteen isolates (0-2 choices) and six neglectees (3 choices), using Bronfenbrenner's criteria for groups of between ten and fifty persons (Bronfenbrenner, 1945; Gronlund, 1959).

TABLE 5

INCIDENCE OF SPEECH DEFECTS IN E.S.N. AND CONTROL SAMPLES.

Group		Total	Pupils with speech defect	
Boys		N	%	$\chi^2$ (d.f.1)
Junior	E.S.N. ....	50	46.0	3.9†
	Controls ....	50	24.0	
Senior	E.S.N. ....	60	40.0	9.0*
	Controls ....	60	15.0	
Total	E.S.N. ....	110	42.7	12.07*
	Controls ....	110	19.1	
Girls				
Junior	E.S.N. ....	23	21.8	n.s.
	Controls ...	23	13.1	
Senior	E.S.N. ....	36	36.1	5.79†
	Controls ....	36	8.3	
Total	E.S.N. ....	59	30.5	6.55†
	Controls ....	59	10.2	
Total (B. and G.)	E.S.N. ....	169	38.5	18.28*
	Controls ....	169	16.0	

\* Significant at 1 per cent. level or beyond.

† Significant at 5 per cent. level or beyond.

## VII.—DISCUSSION.

The multiple nature of the handicap of E.S.N. children in special schools is amply illustrated by the findings of this enquiry. Although the teachers felt that 143 out of the 169 E.S.N. children (85 per cent.) were benefiting from the education given at the special school, many of these children presented problems to them arising from emotional and social maladjustment, physical handicap, speech defects, unsatisfactory attendance at school and a poor home background. Of the others, three children were considered not to be benefiting

from special schooling; in the case of fifteen pupils, the teachers thought it doubtful whether any real benefit was being derived from special schooling; and in eight cases, it was considered too early to say whether definite progress was being made. Seven children were said to be adversely affecting the progress of other pupils.

In a number of cases, the teachers' task was exacerbated by a lack of co-operation from the home. The head teachers reported that over a third of the parents were indifferent as far as the child's education was concerned, or even antagonistic to the child being in a special school. Some parents very much resented the child's transfer from an ordinary school, though their attitude sometimes changed when the child settled down happily in a special school.

The teachers frequently commented that individual pupils had made progress in their emotional and social development, even if they had not made marked scholastic progress. Eighty-two (45 per cent. of the sample) were said to have improved in their emotional adjustment since entry. However, the Bristol Guides did not reveal more than a slight decrease in maladjustment amongst the senior boys as compared with the junior boys, and amongst the girls there was a higher proportion of maladjusted pupils in the senior group than in the junior E.S.N. classes. This suggests that maladjusted behaviour in E.S.N. children tends to persist over a long period, or that new sources of disturbance affect the E.S.N. pupil during adolescence.

The high incidence of maladjustment and unsettled behaviour amongst E.S.N. pupils suggests that a more therapeutic approach is needed in school, and that the school needs the support of child guidance services, many of which have so far offered little positive help to maladjusted E.S.N. children. Few facilities exist at present for the treatment of these children. As Evans (1956) has pointed out, it is very difficult to find placement for dull children in schools for the maladjusted and virtually impossible to secure psychiatric or psychological treatment for maladjusted children of subnormal intelligence. Even if individual treatment is not warranted or cannot be given, supportive help and advice from child guidance staff is required by special schools, especially as these schools are often expected to deal with pupils whose behaviour is disturbing in the ordinary school.

The nature of the children's maladjustment indicates that they particularly need positive and secure relationships with adults and protection from strain-producing situations, especially from a feeling of rejection. Stott (1963) has suggested that the types of disturbance-pattern which particularly characterize the present E.S.N. sample (HA/XA, D and U) may indicate a fear of rejection and exposure to strain-producing situations. There is an urgent need for forging stronger links between home and school, and for solving the problem of irregular attendance, both by strengthening the social welfare services available to the school and by the provision of more residential schools for those pupils with extremely poor home conditions. However, the residential school, as the Underwood Report (1955) has pointed out, can do little without co-operation from the home and a plan for rehabilitating the child in the family.

The high incidence of speech defects amongst E.S.N. children clearly points to the need for increased facilities for dealing with this problem, as well as for more research into the E.S.N. child's problems in oral communication. Very specialised help, too, is needed for E.S.N. children with gross physical handicaps, who present a difficult task to a school which is not equipped to meet their needs.



ACKNOWLEDGEMENTS.—The writer wishes to record his indebtedness to, and to thank, the following for their help and co-operation: the Directors of Education and Medical Officers for Breconshire, Cardiff, Carmarthenshire, Cardiganshire, Glamorgan, Monmouthshire, Newport, Pembrokeshire, and Swansea; the Children's Officer for Cardiff; all the head teachers and staff of the E.S.N. and ordinary schools concerned; Miss Elisabeth Gruber, Research Officer to the project; Miss Margaret Lewis, Secretary to the project; Mr. R.R. Dale and Dr. D. H. Stott, for valuable comments on this paper; and the Ministry of Education, for their interest and financial support.

## VIII.—REFERENCES.

- BELFIELD, D. J. (1963). *The Social Adjustment of Most Accepted and Least Accepted Children in Junior Schools*. M.Ed. Thesis, Univ. of Manchester.
- BRONFENBRENNER, U. (1945). *The Measurement of Sociometric Status, Structure and Development*. Sociometry Monographs, No. 6, New York: Beacon House.
- BURT, C. (1937). *The Backward Child*. London: University of London Press.
- CHAZAN, M. (1962). The relationship between maladjustment and backwardness. *Educ. Review*, 15, 1, 54-61.
- CHAZAN, M. (1963, a). Maladjustment, attainment and sociometric status. *University College of Swansea Faculty of Education Journal*, pp. 4-7.
- CHAZAN, M. (1963, b). Maladjusted Pupils: Trends in post-war theory and practice. *Educ Research*, VI, 1, pp. 29-41.
- DOUGLAS, J. W. B., and BLUMFIELD, J. M. (1958). *Children Under Five*. London: George Allen and Unwin.
- DUNCAN, JOHN (1947). *The Education of the Ordinary Child*. London: Nelson.
- DUNN, LLOYD M. (Ed.) (1963). *Exceptional Children in the Schools*. New York: Holt, Rinehart and Winston.
- EVANS, D. (1956). *An Experimental Study of a Group of Seriously-Maladjusted Educationally Subnormal Children*. M.A. (Education) Thesis, University of Birmingham.
- FERGUSON, GEORGE A. (1959). *Statistical Analysis in Psychology and Education*. New York: McGraw-Hill Book Co.
- GRONLUND, N. E. (1959). *Sociometry in the Classroom*. New York: Harper and Bros.
- HUTT, M. L., and GIBBY, R. G. (1958). *The Mentally Retarded Child*. Boston: Allyn and Bacon.
- JOHNSON, G. O. (1950). A study of the social position of mentally handicapped children in the regular grades. *Amer. J. Ment. Defic.*, 55, 60-89.
- JOHNSON, G. O., and KIRK, S. A. (1950). Are mentally handicapped children segregated in the regular grades? *Except. Child.*, 17, 65-68, 87-88.
- MCLACHLAN, D. G. (1955). Emotional aspects of the backward child. *Amer. J. Ment. Defic.*, 60, pp. 323-330.
- MILLER, R. V. (1956). Social status and socioemphatic differences among mentally superior, mentally typical, and mentally retarded children. *Except. Child.*, 23, 114-119.
- MINISTRY OF EDUCATION (1955). *Report of the Committee on Maladjusted Children*. London: H.M.S.O.
- MINISTRY OF EDUCATION (1962). *The Health of the School Child, 1960 and 1961*. London: H.M.S.O.
- MORAN, R. E. (1960). Levels of attainment of educable subnormal adolescents. *Brit. J. Educ. Psychol.*, 30, pp. 201-210.
- PETRIE, I. R. J. (1962). Residential treatment of maladjusted children; a study of some factors related to progress in adjustment. *Brit. J. Educ. Psychol.*, 32, 29-37.
- PRINGLE, M. L. KELLMER, and BOSSIO, V. (1958). A study of deprived children, Part I—Intellectual, emotional and social development. *Vita Humana*, 1, pp. 65-92.
- RUSHTON, C. S., and STOCKWIN, A. E. (1963). Changes in Terman-Merrill I.Q.s. of educationally subnormal boys. *Brit. J. Educ. Psychol.*, 33, pp. 132-141.
- STEIN, ZENA, and SUSSER, M. (1960). Families of dull children, Part 2—Identifying family types and subcultures. *J. Ment. Sci.*, 106, pp. 1296-1303.

- STOTT, D. H. (1960). Observations on retest discrepancy in mentally subnormal children. *Brit. J. Educ. Psychol.*, 30, pp. 211-219.
- STOTT, D. H. (1961). *Delinquency Prediction Instrument*. London: U.L.P.
- STOTT, D. H. (1963). *The Social Adjustment of Children* (2nd edition). U.L.P.
- TANSLEY, A. E. (1951). *The Use of Mental and Scholastic Test Data and Case History Information as Indications for the Treatment of E.S.N. Children*. M.Ed. Thesis. Leeds.
- TANSLEY, A. E., and GULLIFORD, R. (1960). *The Education of Slow Learning Children*. London: Routledge and Kegan Paul.
- VERNON, P. E. (1964). *Personality Assessment*. London: Methuen.
- WALKER, GALE, H. (1950). Social and emotional problems of the mentally retarded child. *Amer. J. Ment. Defic.*, 55, pp. 132-138.

(Manuscript received 16th April, 1964)



# REASONING ABILITY AND PERSONALITY DIFFERENCES AMONG STUDENT-TEACHERS

By E. G. S. EVANS

(*University of Southampton Institute of Education*)

**SUMMARY.** This study compares the reasoning abilities and personality characteristics of two representative samples of student-teachers, namely, university graduates and training college students. Results show that at least 20 per cent. of training college students have the necessary intellectual ability to pursue academic work to degree standard. On tests of extraversion and sociability, the same students were significantly better than the graduates. There was no difference between the two groups in terms of emotional stability and self-confidence. The research suggests that universities should pay more attention to the personal problems of its students whilst training colleges need to consider forms of education most likely to benefit those embarking on advanced academic work.

## I.—INTRODUCTION.

AN important need in education at present is for systematic research into the abilities and personal qualities of student-teachers. This problem has a special urgency in view of the Robbins' recommendations for an immediate extension in the number and size of training colleges as well as in the professional and academic studies of those who attend them. But such proposals, however enlightened they might be, depend for their successful implementation on empirical evidence of the student's ability to profit from the changes envisaged in their courses. The present investigation is directed to this end and by examining the intellectual and personal characteristics of two representative samples of student-teachers, aims to show that such attributes as verbal reasoning ability, mental speed and accuracy, stability, self-confidence and social adjustment may be usefully taken into account by those concerned with the planning of future selection and training procedures.

## II.—OUTLINE OF PREVIOUS RESEARCH.

The psychological assessment of student-teachers has been the subject recently of numerous British and American studies. Of the former, the most recent are those of Allen (1963), Phillips (1963), Herbert and Turnbull (1963) and Warburton, Butcher and Forrest (1963). Allen's work reviews most of the major studies concerned with the professional training of teachers; that of Phillips compares the self-concepts of training college students from two different cultures while the investigations of Herbert and that of Warburton, *et al.*, deal with the measurement of students' personalities in relation to their teaching abilities. Of these enquiries, that of Warburton and his colleagues deserves special mention. It was conducted with a sample of 113 graduates in a university department of education and examined the predictive value of different intelligence and personality tests in terms of three criteria, namely, final teaching mark, final theory mark and award of the graduate certificate. Analysis of results showed that degree class was the best single predictor of theory mark and of final award and that a personality questionnaire was the best predictor of teaching mark. Warburton concluded: "It would, therefore, appear best in selecting potential teachers to concentrate on attainment and general culture



rather than ability, on stability rather than on extraversion or introversion, on academic rather than on aesthetic interests and on participation in social activities rather than on social or domestic background."

Other researches involving different student-categories but concerned with the measurement of personality variables in relation to intelligence and academic achievement have been carried out by Eysenck (1963), Furneaux (1962), Savage (1962), Biggs (1962), Lynn and Gordon (1961) and Lynn (1959). The present study is in line with these and attempts to demonstrate that no simple cause-and-effect interaction exists between reasoning ability, educational achievement and such traits as stability, confidence and social adjustment, but that a combination of many factors rather than any one is invariably responsible for the student's progress.

### III.—PLAN OF THE INVESTIGATION.

(1) *The Sample.* The students who took part in the investigation were drawn from four training colleges and a university department of education and numbered 289 (131 men, 158 women). The training college sample consisted of 145 students (fifty-five men, ninety women) equitably distributed in respect of age, education and year of training. Their ages ranged from 18.4 years to 24.3 years (mean 20.53 years, S.D. 1.41 years), while their educational achievement or six G.C.E. 'O' level passes to one or more 'A' level subjects. The second group of students comprised 144 graduates (seventy-six men, sixty-eight women) who were pursuing a one-year teacher-training course in a university. Their ages extended from 21.0 years to 25.1 years (mean 22.59 years, S.D. 0.96 years). The mean difference in age between the two groups of students was thus just over two years. The graduates' degrees were mostly in arts or science, more than half being honours degrees and the remainder ordinary and general degrees.

(2) *The Test Battery.* Three different kinds of standardised group tests were administered uniformly to both samples and were chosen to give measures of (a) verbal reasoning ability, (b) mental speed and accuracy, and (c) the following personality attributes: neuroticism, extraversion, self-confidence, sociability (and their opposites). A score for educational achievement was also obtained.

(a) *Reasoning Ability.* This was assessed by Heim's A.H.5 and Valentine's Reasoning Tests, the reliability of which, over student groups, is reported to be .84 and .83, respectively. A third measure was provided by the Nufferno Level Test (GL/2C.46), 'level' being defined as "the score that results when an intelligence test is designed to reward persistent application as well as ability" (Furneaux, 1963). Similar in content to the letter-series tests used by Thurstone in his Primary Mental Abilities battery, the Nufferno test claims to discriminate effectively between the higher levels of intelligence. Its accuracy is expressed in terms of standard errors of estimate rather than in terms of reliability or standard error, while its validity, with Anstey's Dominoes Test of 'g' as the criterion, is reported to be .87. Percentile norms are available for ages 18-30 years.

(b) *Mental Speed and Accuracy.* Separate scores for these two aspects of reasoning ability were provided by means of the Nufferno Speed and Accuracy Tests (GIS. 14.E36). Standard error of estimate-data are again given which, under optimum conditions, are said to correspond to reliabilities greater than .90. Percentile norms derived from secondary schools and universities are also provided.



(c) *Academic Achievement.* A measure of academic achievement was obtained from the students' G.C.E. results, one mark and three marks, respectively, being given for each 'O' and 'A' level subject. It must be noted, however, that these marks ignored the relatively high academic qualifications of the graduates and simply represented the basic educational attainments common to all the students.

(d) *Personality Tests.* The Maudsley Personality Inventory provided scores for neuroticism and extraversion while the revised Bernreuter Inventory gave measures of self-confidence and sociability. The Maudsley neuroticism scale aims to discriminate between individuals who are tense, excitable and nervous and those who are relaxed and apparently well-adjusted; the extraversion scale between those who are friendly and co-operative and those who are not. The Bernreuter confidence test differentiates between individuals who are self-assured and those who tend to be handicapped by self-consciousness and feelings of inferiority. In this test low or negative scores indicate confidence and high scores the lack of it. Similarly, on the sociability scale, low scores are associated with good sociability, high scores with solitariness and independence.

(3) *Method of Analysis.* The first part of the enquiry was based on a comparison of the mean scores obtained by each group of students on each test. The scores were then correlated for each group in turn and for the combined sample, differences between coefficients being tested for significance in accordance with the usual criterion. Factors of heterogeneity such as age and education were held constant by partial correlation. An attempt was also made to classify the students into temperamental categories on the basis of their performance on the Maudsley Personality Inventory. Cutting-off points of  $N=25$ ,  $E=25$  were used for this purpose. The scores on the sociability and self-confidence tests were then related to these categories.

#### IV.—STATEMENT OF RESULTS.

(1) *Reasoning Ability.* As may be seen from Table 1, the mean scores obtained by the graduates on Valentine's Reasoning Test and that of Heim were significantly higher than those of the training college students. Whilst this could have been predicted on the basis of the graduate's selection for University at 18+, it should be noted, in fairness to the training college students, that 20 per cent. of them obtained higher scores than the average score of the graduates (Table 2). On a slightly lower standard, namely, that of the graduate's 25th percentile score, 43 per cent. of the training college students did better. In this latter comparison there were significantly more training college women than men.

(2) *Mental Speed and Accuracy.* There were no significant differences between the two groups of students in respect of the speed, level and accuracy of their mental functioning. University women graduates tended to be faster at the letter-series problems than both training college men and university men, but the difference in score just failed to reach the .05 level of significance. Moreover, neither the Nufferno Level Test, which is really another measure of problem-solving ability, nor the Nufferno Accuracy Test discriminated between the students as effectively as the Valentine or Heim tests. But, on the other hand, they correlated well with the two latter tests and, as we shall also see, with certain of the personality measures.

TABLE 1  
MEANS AND STANDARD DEVIATIONS.

Test	N	Training College Students				University Graduates			
		Men		Women		Men		Women	
		Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Valentine Reasoning....	192	27.08	11.71	27.71	12.20	40.18	13.60	36.18	12.74
Heim A.H.5 .....	208	33.74	6.95	35.53	7.94	41.08	8.05	39.14	7.38
Nufferno Level .....	170	20.99	6.64	21.94	6.89	23.64	6.81	22.94	7.22
Nufferno Speed .....	141	22.15	1.12	22.24	1.25	22.06	1.58	22.76	1.26
Nufferno Accuracy .....	141	14.91	2.65	15.40	2.43	15.44	2.15	15.75	2.66
M.P.I. Neuroticism ....	258	25.71	10.70	25.43	9.87	25.61	12.54	26.97	9.16
M.P.I. Extraversion ....	258	28.71	8.38	29.37	9.26	22.88	9.84	25.56	9.60
Bernreuter Confidence*	184	-7.17	110.25	16.93	77.02	9.48	101.17	30.34	75.66
Bernreuter Sociability*	184	-2.63	64.81	-29.97	57.92	15.24	52.12	-8.66	53.19

\* The size of the S.D.'s on the Bernreuter tests is due to the scores ranging from -200 to +245.

TABLE 2  
PERCENTAGE OF TRAINING COLLEGE STUDENTS WITH REASONING ABILITY SCORES GREATER THAN THE MEAN AND QUARTILE SCORES OF UNIVERSITY GRADUATES.

Test	N	Men		Women		Overall Percentages	
		% > X	% > Q <sub>1</sub>	% > X	% > Q <sub>1</sub>	% > X	% > Q <sub>1</sub>
Valentine .....	117 132	17.31	36.54	18.46	44.61	17.95	40.58
Heim A.H.5. ....		19.57	41.30	21.74	51.16	21.97	46.23
Overall Percentages .....		18.44	38.93	20.10	47.88	19.96	43.40
S.E.'s % .....		2.45	3.08	2.83	3.16	2.82	3.13

(3) *Personality Differences.* Significant differences were obtained on the extraversion and sociability scales with training college students gaining higher scores than the graduates. As a result both training college men and women appeared to be much more extraverted than university men. The mean extraversion score of the university women, however, was not significantly different from that of the training college men. With this exception, the results are consistent with the findings of Furneaux (1957) and Lynn (1959) insofar as the detrimental effects of extraversion seem to manifest themselves most clearly below university level. On the sociability test, training college women obtained better scores than university men and, for that matter, training college men. But no difference in sociability was apparent between training college women and their university counterparts. The neuroticism and self-confidence scales likewise produced no differences between any of the groups though there was a tendency for training college men to be more confident than university women.



(4) *Temperamental Categories.* When the students were divided into the categories of neurotic or stable and introverted or extraverted, on the basis of their scores on the Maudsley Personality Inventory, the graduates tended to be equitably distributed over all four divisions whereas the training college students came firmly within one or other of the extravert categories (Table 3). These dichotomies, however, are only approximate and have been included merely to indicate possible differences in temperament between the two student-groups.

TABLE 3  
CLASSIFICATION OF STUDENTS ACCORDING TO TEMPERAMENT.

Temperamental Category .....	T.C. students		Graduates		Total	
	N	%	N	%	N	%
Stable Extravert .....	41	36.9	19	27.9	60	33.5
Neurotic Extravert .....	35	31.5	13	19.1	48	26.8
Neurotic Introvert .....	19	17.1	21	30.8	40	22.3
Stable Introvert .....	16	14.4	15	22.0	31	17.3

The scores for confidence and sociability for the combined sample were then considered in relation to the above classification and, as may be seen from Table 4, stable extraverts were significantly more confident, and neurotic introverts significantly less confident than any other category. There was no difference in confidence between stable introverts and neurotic extraverts. On the sociability test, both stable and neurotic extraverts did significantly better than the neurotic introverts, but between the other categories no difference in sociability was apparent.

TABLE 4  
MEAN SCORES FOR CONFIDENCE AND SOCIABILITY IN RELATION TO TEMPERAMENTAL CATEGORIES.

Temperamental Category	N	Confidence		Sociability	
		Mean	S.D.	Mean	S.D.
Stable Extraverts .....	60	-57.21	61.35	-29.18	81.06
Neurotic Extraverts .....	48	32.70	67.71	-24.08	56.01
Neurotic Introverts .....	40	103.15	77.54	31.25	52.98
Stable Introverts .....	31	2.26	84.28	0.90	60.76

(5) *The Intercorrelations.* 70 per cent. of the correlation coefficients were significant, the great majority reaching the .01 level or better (Table 5). All the reasoning tests, including the Nufferno Level Test, correlated well with one another, as did the Nufferno Speed and Accuracy Test. So far as the two latter measures were concerned, accuracy correlated much better with reasoning ability than did speed. But speed correlated positively with confidence and negatively with neuroticism.

TABLE 5  
INTER-CORRELATIONS.

Correlated Variables	Training College	University Graduates	Combined Sample
Valentine R : Heim A.H.5. ....	.408	.502	.473
Valentine R : Nufferno Level .....	.343	.306	.362
Valentine R : Nufferno Accuracy ....	.296	.313	.305
Valentine R : Nufferno Speed .....	.115	.263	.169
Heim A.H.5 : Nufferno Level .....	.543	.637	.584
Heim A.H.5. : Nufferno Accuracy ....	.508	.471	.496
Heim A.H.5. : Nufferno Speed .....	.319	.335	.330
Nufferno Level : Nufferno Accuracy ..	.416	.576	.456
Nufferno Level : Nufferno Speed .....	.156	.249	.190
Neuroticism : Valentine R. ....	-.190	-.442	-.234
Neuroticism : Heim A.H.5. ....	-.174	-.343	-.205
Neuroticism : Nufferno Speed .....	-.116	-.495	-.219
Neuroticism : Lack of Confidence ...	.643	.627	.638
Neuroticism : Unsociability .....	.203	.239	.215
Confidence : Heim A.H.5. ....	.011	.466	.105
Confidence : Valentine R. ....	.092	.436	.139
Confidence : Nufferno Level .....	-.131	.392	-.020
Confidence : Nufferno Speed .....	.053	.491	.167
Extraversion : Heim A.H.5. ....	-.117	.389	-.030
Extraversion : Nufferno Accuracy ....	-.099	.429	.125
Extraversion : Confidence .....	.451	.611	.506
Extraversion : Sociability .....	.296	.605	.403
P (.05) .....	.188	.339	.165
(.01) .....	.246	.436	.216

The personality tests inter-correlated satisfactorily in all three samples, the highest relationships being those for neuroticism and lack of confidence (.643), extraversion and confidence (.611) and extraversion and sociability (.605).

The relation of these personality tests to the reasoning tests was also most marked. Thus, the Valentine and Heim reasoning tests, like the Nufferno Speed Test, correlated positively with confidence and negatively with neuroticism. The Nufferno Level Test was also found to be significantly related to confidence. But these relationships, it should be noted, existed only in the case of the graduates and not with the training college sample. In the latter group, and with the single exception of a significant negative correlation between neuroticism and Valentine's test, no relationships of any note occurred between reasoning ability and such traits as confidence, extraversion and sociability. Furthermore, when the coefficients obtained for the two samples in respect of these tests were transformed into Fisher's  $z$ 's and the differences tested, significant 't' ratios, mostly at the .01 level, were secured in all but two cases (Table 6).



TABLE 6  
SIGNIFICANT DIFFERENCES BETWEEN COEFFICIENTS.

Correlated Variables		Training College Students		Graduates		S.E.D	t
		Z	S.E. <sub>z</sub>	Z	S.E. <sub>z</sub>		
Neuroticism	: Heim A.H.5	-.176	.094	-.358	.157	.183	0.99
Neuroticism	: Valentine R ..	-.192	.093	-.475	.139	.167	1.69
Neuroticism	: Nufferno Speed	-.117	.095	-.543	.127	.159	2.69
Confidence	: Valentine R.	.092	.096	.467	.140	.170	2.25
Confidence	: Heim A.H.5	.011	.097	.505	.134	.175	2.99
Confidence	: Nufferno Level	-.132	.095	.402	.151	.178	3.00
Confidence	: Nufferno Speed	.053	.096	.537	.128	.160	3.03
Extraversion	: Nufferno Accuracy ....	-.099	.096	.459	.142	.171	3.26
Extraversion	: Sociability ...	.305	.087	.709	.087	.123	3.28
Extraversion	: Heim A.H.5.	-.118	.010	.416	.149	.149	3.57

N = 140    P(.01) = 2.61, P(.05) = 1.98.

In this matter, the coefficients calculated for the graduate's sample were invariably higher than those of the training college students. The reasons for this are discussed later.

(6) *The Influence of Age and Education.* Taking the training college students and university graduates together, chronological age was found to correlate significantly with academic achievement (.425), neuroticism (.171), extraversion (-.201) and mental accuracy (-.160). Academic achievement, in its turn, correlated with Valentine (.240), Heim (.215) and extraversion (-.317).

TABLE 7  
PARTIAL COEFFICIENTS OF CORRELATION FOR REASONING ABILITY, ACCURACY AND SPEED.

A.—With Chronological Age Constant.					
Test	Valentine	Heim	Nufferno	Accuracy	Speed
Valentine ..	—	—	—	—	—
Heim A.H.5.	.484	.594	—	—	—
Nufferno ...	.379	.513	.454	.448	—
Accuracy ...	.334	.335	.185	—	—
Speed .....	.185				
B.—With Academic Achievement Constant.					
Test	Valentine	Heim	Nufferno	Accuracy	Speed
Valentine ..	—	—	—	—	—
Heim A.H.5.	.447	.582	—	—	—
Nufferno ...	.348	.527	.471	.459	—
Accuracy ...	.351	.351	.197	—	—
Speed .....	.183				

N = 140.    P(.05) = .165    P(.01) = .216.

Holding age constant in the combined sample produced no substantial changes in any of the coefficients for reasoning ability and personality, as Tables 7 and 8 show. But when academic achievement was partialled out, the relationship between age and extraversion and that between age and accuracy ceased to be significant.

TABLE 8

PARTIAL COEFFICIENTS OF CORRELATION FOR PERSONALITY VARIABLES.

Variable	With C.A. constant		With A.A. constant	
	Neuroticism	Extraversion	Neuroticism	Extraversion
Confidence .....	-.638	.505	-.718	.511
Sociability .....	-.195	.390	-.233	.414

N=140.  $P(.05) = .165$ .  $P(.01) = .216$

#### V.—DISCUSSION AND CONCLUSIONS.

Three points in the above findings call for discussion: (1) the percentage of training college students who obtained higher scores than the graduates on the tests of reasoning ability; (2) the better performance of the same students on the extraversion, sociability and confidence tests and (3) the relation of the personality variables to reasoning ability and academic success.

(1) That so many training college students, women particularly, secured better scores than the graduates on the reasoning tests, and were just as proficient as the latter on the speed and accuracy tests, raises the question as to why they did not gain higher academic distinctions at school and are not now pursuing university degree courses. The answer probably lies in their different interests and experiences at school and in the fact that on leaving, they chose the cultural and social amenities of training college life to what can often be an arduous intellectual training at university degree level. Even more, perhaps, they prefer dealing with people and social situations to working with things or ideas. But, whatever other factors of temperament or background might have influenced them at school and in their subsequent choice of a teaching career, one thing is crystal clear. Insofar as university work depends on intelligence, it is recognised, however, that the form of higher education likely to be most beneficial to such students has still to be developed and 'tailored' to their special abilities and interests. In particular, the structure of Ed.B. degree courses, their subjects and syllabuses, will have to be made very flexible to allow for the students' differing needs. Another important research topic related to the problem of providing appropriate degree courses would be to investigate those students of high intelligence who are under-functioning because of insecurity, lack of confidence, loneliness and other difficulties but who, given the right kind of guidance, could cope satisfactorily with advanced academic work.

\* This finding is consistent with the Robbins' estimate of 25 per cent. of such students who might take degree courses if facilities were available.



(2) The personality tests showed that, while university graduates were just as stable as training college students, they were inclined to be more introverted. Furthermore, in spite of the differences in age and intelligence between the two groups, university men were less sociable than training college women, and university women less confident than training college men. Again, one cannot help wondering whether these differences are induced by the different experiences the students have in university and training college. One would expect a greater social maturity and confidence to develop and, indeed, be accelerated as a function of a university education, but this does not seem to happen. On the contrary, the present findings suggest that though many students go through a university and gain a fair level of academic success and competence in their special subject fields, they do not appear to achieve much in terms of their personal development. For example, 30.88 per cent. of the graduates came within the category of neurotic introvert as compared with 17.12 per cent. of training college students (Table 3).<sup>\*</sup> This fact, together with the poor scores obtained by many graduates on the confidence and sociability tests suggest that universities could learn a thing or two from training colleges in their approach to students' needs and problems and by so doing, bring about a radical change of attitude towards the purpose of 'reading for a degree' and towards the purpose of education in general.<sup>†</sup>

(3) Some comment is also necessary concerning the relation of the personality variables, particularly neuroticism and extraversion, to academic attainment. In the present study no significant correlation was obtained between neuroticism and scholastic success. This may have been due to the arbitrary nature of the attainment measures, for the mean neuroticism scores of training college students and graduates alike were similar to those obtained by Furneaux (1957), Lynn (1959), Lynn and Gordon (1961), Savage (1962) and Biggs (1962) and were consistent with the hypothesis that a certain amount of anxiety and personal tension are necessary for academic success. On the other hand, Warburton's (1964) review shows that twenty-nine out of forty-two recent researches claim that it is stability rather than anxiety that exerts the greater influence on attainment. The overall results are thus indecisive and suggest that further research into this problem is needed.

The position is much clearer so far as extraversion and scholastic success are concerned, for the present study corroborates the findings of Furneaux (1957), Lynn (1960), Cattell (1960), Holmes (1960) and Savage (1962), in showing that introversion, and not extraversion, is the positive influence on attainment. Warburton mentions eighteen out of twenty studies as reporting similar results.

(4) Finally, in the matter of the relationship between reasoning ability and the personality variables, it was noted that the correlation coefficients for the graduate's sample were significantly higher than those of the training college students. In descriptive terms, the reason for the differences is that whereas with the graduates high scores on intelligence, speed and accuracy were

<sup>\*</sup> As one student has written recently: "Anyone who has come into close contact with undergraduates must be struck by the sizeable proportion who seem ill-adjusted, bored and frustrated. For many of the most intelligent and imaginative students, university has apparently not provided the intellectual stimulus and all-round education that it claims. Rather is it a place where 'pebbles are polished and diamonds are dimmed'."—"Youth—Humbug," *The Sunday Times*, October 13th, 1963.

<sup>†</sup> Another writer, Zweig (1963) states: "The social revolution of university education has produced new problems. It has introduced a more harassed, more anxious and more worried type of student, and a more harassed atmosphere at the university... What is needed is more spontaneity, more lightness, more colour, more warmth, more individual care, more wit and humour. Where is the student who sings *Gaudeamus igitur, iuvenes dum sumus?*"



accompanied by high scores for stability, confidence, sociability and so on, these same personal qualities tended to be equitably distributed throughout the whole range of ability in the training college sample. The most intelligent training college students, for example, were no more confident, sociable or extraverted than those students of average ability who, in their turn, were no better than those of below average intelligence. This, then, explains the lack of correlation between reasoning ability and the personality traits in the training college sample.

In considering possible reasons for the differences in the correlation coefficients of the two samples one is again forced to ask whether they were, brought about by virtue of the different lives the students lead in training college and university, and especially by the different social and academic experiences they have in these places. As a result, training college students, irrespective of their intellectual abilities, appear relatively happy and settled in their work whilst many university graduates whose intelligence is being misdirected are inclined to be introverted, solitary and unsure of themselves. That well-adjusted personalities are not always the prerogative of the highly intelligent may be inferred from a comparison of the mean personality scores obtained by the two groups (Tables 1 and 4). Once again the lesson is clear and, as Drever (1963) says: "Universities should learn from the training colleges to give professional and directed education . . . they need machinery that translates the results of educational research into changes of syllabus and organisation."

#### VI.—REFERENCES.

- <sup>1</sup> ALLEN, ERIC A. (1963). Professional training of teachers—A review of research. *Educ. Res.*, 5, 3, June.
- <sup>2</sup> BIGGS, J. R. (1962). The relation of neuroticism and extraversion to intelligence and educational achievement. *Brit. J. Educ. Psychol.*, 32, 2, June.
- <sup>3</sup> CATTELL, R. B. (n.d.). Predicting success in the academic aspects of military training schools. *I.P.A.T. Information Bulletin*, No. 4, Department of Psychology, University of Illinois.
- <sup>4</sup> DREVER, J. (1963). *Times Educational Supplement*, p. 883, Dec. 13th.
- <sup>5</sup> EYSENCK, H. J. (1963). On the relation of neuroticism and extraversion to intelligence and educational achievement. *Brit. J. Educ. Psychol.*, 33, June.
- <sup>6</sup> FURNEAUX, W. D. (1957). *Student Selection*. Report to Imperial College of Science and Technology.
- <sup>7</sup> FURNEAUX, W. D. (1962). The psychologist and the university. *Universities Quarterly*, 17, 1 Dec.
- <sup>8</sup> HERBERT, N., and TURNBULL, G. H. (1963). Personality factors and effective progress in teaching. *Educ. Rev.*, 16, 1, Nov.
- <sup>9</sup> HOLMES, F. J. (n.d.). Predicting academic success on a general college curriculum. *I.P.A.T. Information Bulletin*, No. 4, Department of Psychology, University of Illinois.
- <sup>10</sup> LYNN, R. (1959). Two personality characteristics related to academic achievement. *Brit. J. Educ. Psychol.*, 29, 3, Nov.
- <sup>11</sup> LYNN, R., and GORDON, I. E. (1961). The relation of neuroticism and extraversion to intelligence and educational achievement. *Brit. J. Educ. Psychol.*, 31, 2, June.
- <sup>12</sup> PHILLIPS, A. S. (1963). The self concepts of trainee teachers in two sub-cultures. *Brit. J. Educ. Psychol.*, 33, June.
- <sup>13</sup> SAVAGE, R. D. (1962). Personality factors and academic achievement. *Brit. J. Educ. Psychol.*, 32, 3 Nov.
- <sup>14</sup> WARBURTON, F., BUTCHER, H. J., and FORREST, G. (1963). Predicting student performance in a university department of education. *Brit. J. Educ. Psychol.*, 33, 1, Feb.
- <sup>15</sup> WARBURTON, F. W. (n.d.). *A Survey of Researches into the Relationship between Scores in Personality Questionnaires and Scholastic Attainment*. Faculty of Education, University of Manchester, n.d.
- <sup>16</sup> ZWEIG, F. (1963). *The Student in the Age of Anxiety—A Survey of Oxford and Manchester Students*. London: Heinemann.

(Manuscript received 7th May, 1964)



# ZERO ERROR AND PRACTICE EFFECTS IN MORAY HOUSE ENGLISH QUOTIENTS

BY JOHN MACNAMARA

(*St. Patrick's Training College, Dublin*)\*

SUMMARY. Ninety-six children in two Dublin primary schools were each set six tests of the Moray House English Test series. The purpose of the experiment was twofold: (a) to discover whether Moray House English Tests vary in difficulty level and to measure the variation, or zero error, if it exists; (b) to measure the extent to which Moray House English Quotients are influenced by practice. The particular interest of making measurement (b) in Dublin was that the children were, to begin with, totally unfamiliar with such tests whereas for many years British children of the same age cannot be considered to have been so. However, gains with practice were not greater in this than in previously reported investigations.

The existence of zero error in Moray House English Tests was established. The finding is discussed and related to similar findings with other tests, and an attempt is made to discover the reasons for the phenomenon.

## I.—INTRODUCTION.

PILLINER, *et al.* (1960) studied variation in the performance of samples of children on whose work Moray House Verbal Reasoning Tests (M.H.Ts.) have been standardised. They discovered that performance has improved since the late 1940's to such an extent that about 6 quotient points must be added to I.Qs. obtained with recent M.H.Ts. to make them comparable with I.Qs. obtained at the present time with M.H.T. 40 (standardised in 1948). This bias in quotients the author's name 'zero error.' And they attribute it in the case of M.H.Ts. to test-sophistication—"there can be little doubt that this rise in performance level reflects mainly the effect of test-sophistication." In concluding the article, they point out that many research workers, for example many of those who have investigated coaching and practice effects by administering series of tests, have not been sufficiently aware of this particular source of error.

This article reports the findings of an investigation which, though undertaken chiefly to measure practice effects, throws light also on zero error. It differs from the above mentioned study in that the tests employed were Moray House English Tests (M.H.Es.). And because it was conducted on Irish children who were to begin with totally unfamiliar with 'objective' tests of any sort, it differs from most studies of practice effects in Britain. Presumably, all or practically all, the children who took part in such investigations in Britain were test wise to some extent at the beginning.

## II.—THE EXPERIMENTAL INVESTIGATION.

The tests chosen were M.H.E. 14, 18, 23, 25, 29, and 32. The norms for these tests were based on data obtained during the following years, respectively: 1941-44, 1945-48, 1951-53, 1953-55, 1956-58, and 1958-60.

\* I am deeply indebted to Dr. T. Renshaw of the University of Edinburgh Department of Education for assistance with the design of the experiment recorded in this article and with the interpretation of the findings.



The experiment was conducted in two Dublin national schools, one boys' and one girls', which in the local inspector's opinion, were academically above average for Ireland. As a national survey had been carried out with M.H.E. 14 (Macnamara, 1963), Irish norms for that test were available and by referring to them the writer was able to confirm the inspectors' judgment. The children tested in the present investigation were chosen at random from the fifth standard children in their schools, forty-eight boys and forty-eight girls in all, who sat two tests each week for three weeks. If a child was absent at the time when the others were tested, he sat the test when he next came to school. In each school the children were divided at random into six groups of eight each; and in each school each group of eight sat the six tests in a different order from the other groups in the same school. In short, the experiment took the form of a 'latin square' in each school.

In order that their interest might be maintained throughout the experiment, the children in each school were told that they were competing with those of another school, and that they would all be told the results. They were also told that they would hear how they compared with children in Britain, and this aroused considerable interest.

### III.—EXPERIMENTAL RESULTS.

The range of ages in an Irish class or standard is larger than in a British one. The ninety-six children who were tested ranged in age at the beginning of the experiment from 10 years 0 months to 13 years 6 months. This made it difficult to determine the correct E.Q. to assign to certain raw scores, because M.H.E. norms cover only the range from 10 years 0 months to 12 years 0 months. However, the test constructors state that extrapolation up to 12 years 6 months is permissible, and only six of the children (who were randomly distributed among sub-groups) were above that age. The difficulty in the case of those six was likewise met by extrapolating from the table of norms, but in doing so, the age allowance was progressively reduced with age above 12 years 6 months. The reason for the reduction is that tests constructed for use with children of a particular age range are generally too simple for older children. As a result, they do not distinguish between the varying abilities of older children as finely as they do between those of younger children. Consequently, the increase in score per month is less for older than for younger children.

By the procedure outlined above, 576 E.Qs. were obtained. In analysing these statistically, the method given in Lindquist (1953, pp. 285-8), was employed which entails two separate analyses of the 'total' sum of squares (which are outlined in Table 1) under the headings 'between Ss.' and 'within Ss.' The components of the former were derived without reference to the fact that each child took several tests, while in deriving those of the latter, this fact is recognised. It should be noted that differences between groups, and the many interactions between groups and the three main components mentioned in Table 1, are confounded with components which are associated with Tests or Occasions or both.

The appropriate term to use in testing the mean squares in each section of Table I is the mean square designated 'error' in that section. Since none of the ratios between error and interaction terms is significant, and since the mean square for schools is not significant, being less than unity, it is reasonable to disregard fluctuations which are associated with these sources. Two mean



squares, however, are significant, namely those designated 'Occasions' and 'Tests.' Each calls for comment.

In testing the significance of the Occasions mean square, we test the hypothesis that performance level is unaffected by practice. Because this mean square is highly significant ( $F \geq 4.6$  is significant at the 1 per cent. level of probability where there are 5 and 420 degrees of freedom) we reject the hypothesis: The means, which are given in Table 2, reveal the effect of practice.

TABLE 1  
ANALYSIS OF VARIANCE OF E.Qs.

Source.....	DF	SS	MS	F
Between Ss. ....	95	64956	—	—
Schools .....	1	154	154	—
Occasions† × Tests (b).....	5	5962	1192	1.759 (DF=5, 84)
Occasions × Tests × Schools (b)	5	1917	383.4	—
Error (b) .....	84	56923	677.7	—
Within Ss. ....	480	8197	—	—
Occasions .....	5	2389	477.8	5.049** (DF=5, 420)
Tests .....	5	1099	219.8	2.322* (DF=5, 420)
Occasions × Tests (w) .....	20	189	9.45	—
Occasions × Schools .....	5	342	68.4	—
Tests × Schools .....	5	19	3.8	—
Occasions × Tests × Schools (w)	20	184	9.2	—
Error (w) .....	420	3975	94.64	—
Total .....	575	73154	—	—

† By 'Occasion' is meant a period of testing, of which there were six in each school.

TABLE 2  
MEAN E.Q. OBTAINED ON EACH OCCASION.

Occasion .....	1	2	3	4	5	6
Mean E.G. ....	90.09	92.70	94.31	94.51	95.76	96.13

There is a mean increase of about 6 points from the first to the sixth Occasion which may be attributed to practice.

Turning now to the second significant mean square in Table 1 ( $F \geq 2.23$  is significant at the 5 per cent. level of probability where there are 5 and 420 degrees of probability) we reject the hypothesis that mean E.Q. is unrelated to

the test with which it is obtained. Table 3 gives the mean obtained with each of the six tests.

TABLE 3  
M.H.E. TEST MEANS.

Test .....	M.H.E. 14	M.H.E. 18	M.H.E. 23	M.H.E. 25	M.H.E. 29	M.H.E. 32
Norms Compiled ..	1941-44	1945-48	1951-53	1953-55	1956-58	1958-60
Mean E.Q. ....	94.50	95.97	94.28	94.50	92.19	92.06

These means reveal an increase in difficulty level from M.H.E. 14 to M.H.E. 32, resulting in a drop of about 2.5 quotient points. This may be taken as reliable evidence of a change in the performance level of British children over the past twenty years or so.

18 — Discussion.

The present findings relating to the effect of practice are in remarkably close agreement with the findings of similar studies in Britain carried out with verbal reasoning tests. Professor Vernon (1960, p. 131) sums up the latter with the words: "A single practice test gives [a gain of] about 3-4 points, and several practices 5-6 points." The quotation is interesting for two reasons: (i) the greatest gain to be expected from practice alone is about 6 points; (ii) gains are greatest early in the series of tests and they diminish as the series progresses. The effect of practice on E.Qs. has not been studied to the same extent, but the findings of Watts, *et al.* (1952, p. 37) who administered ten complete M.H.Ts. and ten complete M.H.Es. to groups of London children indicate that the effects of coaching and practice on means obtained with the two types of test are almost equal. It is surprising that the present findings should parallel so accurately those of similar studies in Britain seeing that the Irish children at the beginning of the experiment were, one supposed, less familiar than British children with M.H.E.-type questions. However, it does appear unlikely that the Irish children tested would have gained much more than 6 points of E.Q. by extending the number of practice tests beyond six—though neither in Britain nor in Ireland have experiments been extended to include a sufficient number of tests to establish this.

The point has already been made that the children tested were above average for Ireland. In fact, their mean on the first Occasion exceeds the mean obtained in the Irish national survey by about 10 E.Q. points.\* Yet, the means given above in Table 2 all fall below those obtained by the British children on whom the tests were standardised. If we assume, as we must, that British and Irish children do not differ markedly in 'intelligence,' the latter's poor English scores may at first sight be puzzling. However, the present writer has shown elsewhere (Macnamara 1963, chap. 8) that almost the entire difference in mean E.Q. can be traced to two major sources: (1) difference in test sophistication and (2) difference in the amount of school time devoted to English in the two countries. A comparison of figures for Irish national schools with those given by Morris (1959) for Kent show, if the latter are fairly representative of British

\* See Macnamara (1963), pp. 294 sq.



schools generally, that most British schools devote more than twice as much time as Irish schools to teaching English. Thus, there is nothing surprising in the finding that brighter than average Irish children obtained E.Qs. lower than the British average.

These considerations raise the issue of how valid the present findings are for children in general. There is *fairly firm evidence from research in Britain* (Vernon, 1951; Peel 1952 and 1953; Watts *et al.* 1952, Appendix III) that brighter children gain more than duller ones from coaching and practice with verbal reasoning tests. The evidence we have seen suggests that the same is likely to be true of English tests. Thus, it is highly probable that the Dublin children who were tested gained slightly more from practice than a more representative sample of Irish children would have gained. Thus, since the present findings parallel comparable findings in Britain so closely, we may reasonably conclude that work on the effects of coaching and practice in Britain has not, contrary to what one would have expected, been much disturbed by the initial degree of test sophistication among the children who took part in such experiments.

When we turn to zero error, the parallel between M.H.Ts. and M.H.Es. falls down because zero error in M.H.Es. appears to be very much smaller than in M.H.Ts., about 2.5 points of E.Q. as compared with about 6 points of I.Q. A further difference is that zero error enters the M.H.T. series earlier than the M.H.E. one. The first large zero error in M.H.Ts. relative to M.H.T. 40, and probably the first significant one, is in M.H.T. 49, whose standardisation date is 1953; whereas it enters the English series later than M.H.E. 25 which was standardised on data obtained in the years 1953-55.

If we take the view that zero error in both series is to be attributed to the cause indicated by Pilliner, *et al.* (1960), i.e., test sophistication, we might argue that the difference in size of error is due to a difference in level of sophistication relative to the two types of test in the late 1940's and early 1950's. Vernon (Ed. 1957, p. 57) cites evidence that in those years children were more thoroughly coached for English than for verbal reasoning tests. As the effects of coaching and practice are limited, we would then expect an all round increase in coaching and practice to have a greater effect on M.H.T. than on M.H.E. norms. As regards time of appearance, we might argue that the increase in coaching and practice which followed the public controversy on the matter (Vernon 1952), and clearly affected the M.H.T. series almost immediately, did not begin to affect the M.H.E. series until some years later. However, the unsure nature of this argument prompts us to cast about for an alternative explanation.

One alternative explanation is to be found in two pamphlets published by the Ministry of Education (1950 and 1957) on standards of reading in primary schools. The first indicates that during the war there was a fall in standard equivalent to about one year of reading age. The second indicates that by 1956, about half of that loss had been regained. Thus, we might argue that zero error in the tests we have been discussing is due in part at least to solid improvement in the standard of English, as distinct from increased test sophistication. Clearly, both M.H.Ts. and M.H.Es. measure reading comprehension to a very considerable extent. However, it is just as difficult to distinguish the effects of test sophistication and improvement in the standard of English on the ministry's reading test as upon M.H.Ts. or M.H.Es. Moreover, the date of the Ministry's survey, 1956, makes it difficult to interpret the improved performance in both M.H.Es. and the Ministry's test as the result of a single cause.

To sum up them, our findings support those of previous studies which indicate that gains with practice on M.H.Es. are of about the same magnitude and occur in much the same manner as on M.H.Ts. Thus, it is extremely likely that M.H.E., like M.H.T., norms are subject to practice and coaching effects. Furthermore, it has been established that M.H.E. norms reflect a rise in standard of performance over the past fifteen years, though that rise is very much smaller than the one in M.H.T. norms. But, while increased coaching and practice could explain the rise in standard of performance on M.H.Es., the rise might just as easily be due, in whole or in part, to other causes such as an improvement in the standard of English reading.

#### V.—REFERENCES.

- LINDQUIST, E. F. (1953). *Design and Analysis of Experiments in Psychology and Education*. Boston: Houghton Mifflin.
- MACNAMARA, JOHN (1963). *The Use of Irish in Teaching Children from English-Speaking Homes: A Survey of Irish National Schools*. Unpublished Ph.D. Thesis, University of Edinburgh.
- MINISTRY OF EDUCATION (1950). *Reading Ability*. Pamphlet No. 18. London: H.M.S.O.
- MINISTRY OF EDUCATION (1957). *Standards of Reading, 1948-1956*. Pamphlet No. 32. London: H.M.S.O.
- PEEL, E. A. (1952). Practice effects between three consecutive tests of intelligence. *Brit. J. Educ. Psychol.*, 22, 196-199.
- PEEL, E. A. (1953). Footnote on practice effects in intelligence tests. *Brit. J. Educ. Psychol.*, 23, 126.
- PILLINER, A. E. G., SUTHERLAND, J., and TAYLOR, E. G. (1960). Zero error in Moray House verbal reasoning tests. *Brit. J. Educ. Psychol.*, 30, 53-62.
- VERNON, P. E. (1951). Recent investigations of intelligence. *Eugen. Rev.* 43, 125-137.
- VERNON, P. E. (1952). *Intelligence Testing*. *Times Educational Supplement* reprint of articles and letters on practice and coaching.
- VERNON, P. E. (Ed.) (1957). *Secondary School Selection*. London: Methuen.
- VERNON, P. E. (1960). *Intelligence and Attainment Tests*. London: University of London Press.
- WATTS, A. F., PIDGEON, D. A., and YATES, A. (1952). *Secondary School Entrance Examinations*. London: Newnes.

(Manuscript received, 1st April, 1964).



## RESEARCH NOTE

### SOCIAL CLASS DIFFERENTIALS IN VOCABULARY EXPANSION

By GUSTAV JAHODA

(*University of Strathclyde*)

Socio-economic status differences in performance on intelligence and other cognitive tests are among the most firmly established generalisations in psychological research, though their causal interpretation remains a focus of lively controversy (Anastasi, 1960). There is a considerable weight of evidence indicating that the verbal inferiority of working class children may account in part for the differences (Bernstein, 1961). The existence of a verbal handicap is unquestioned, but its fate over time remains somewhat doubtful. Thus, Eells, *et al.* (1951) held that status differences remain constant in magnitude, invoking the 'levelling effect of a common school environment' as one of the reasons. As against this, Bernstein (1961) contends that with increasing age, working class children tend to fall further behind. A study by Ravenette and Kahn (1962) supports Bernstein's view, but suffers from the limitation of being confined to a comparison of verbal and performance I.Qs. within a disturbed working class population. The aim of the present study was to undertake a direct comparison of working and middle class children at two age levels.

#### SUBJECTS AND METHODS.

Since it is the children of semi- and unskilled working class families who are most handicapped linguistically, it was decided to concentrate on these and contrast them with middle class children. Hence schools were selected in which the overwhelming majority of the children came from either one or the other type of background, and on the basis of the information available from school records the sons of skilled manual workers and artisans were eliminated. The drawback is, of course, that the independent variable in this context is social class conceived as home and school environment in combination, so that the independent contribution of the latter cannot be isolated. It would be difficult to find any alternative in a small-scale study, as practically no schools contain sizeable proportions of children from both the lower working class and the central range of the middle class. The schools were, therefore, chosen in areas of maximal social homogeneity, and the primary schools were mostly 'feeders' for the secondary schools, so as to ensure a closely similar type of population at both levels.

In the case of the working-class (*W*) children, no selection was involved between primary and secondary school, but allocation to streams in the secondary school was on the basis of ability; all streams within the appropriate age-ranges were sampled. The middle-class (*M*) schools differed: one had a kind of 'test' at age 5, followed again by differential streaming at the secondary level; in two others there was selective entry after primary school, but the bulk of candidates (whether accepted or rejected) came also from middle-class families. The general aim was thus to achieve a relatively 'pure' representation of lower working and relatively well-do middle class, by avoiding mixed areas where differential educational mobility might be a source of contamination.

The subjects in the first stage consisted of 214 *M* and 151 *W* boys, age ranges being 9½—10½ in the primary, and 13½—14½ in the secondary schools. Mean ages of *M* boys were precisely 10 and 14, those of *W* boys one month higher.

MacArthur and Elley (1963) demonstrated the relative insensitivity of Progressive Matrices, the non-verbal test used, to social status differences: the younger boys were given the Coloured and the older ones the Standard Progressive Matrices.

For the vocabulary assessment, the Self-Judging Vocabulary Scale (S.J.-Scale) devised by Heim and Watts (1961) was used, form *aA* at age 10 and form *AE* at age 14. Sub-test *A* was common to both, whilst *a* was easier and *E* more difficult than *A*. Unlimited time was allowed, and the procedure usually took 2-3 hours, including intervals.

On the basis of matrices scores, matched groups of *W* and *M* boys were established, consisting of thirty-four pairs at age 10 and twenty-six at age 14; the overlap was relatively small, resulting in a sharp reduction from the numbers tested in the first stage.

#### RESULTS.

Only the crucial comparison relating to groups matched for non-verbal intelligence will be presented. A single measure was derived for each pair, by expressing the Self-Judging Vocabulary (S.J.) score of each working-class boy as a proportion of that of his middle-class partner. The Mann-Whitney U test was applied to the resulting distribution. The outcome appears in Table 1, which consists of two parts.

TABLE 1

MEAN AND PERCENTAGE SCORES \* BY AGE, SOCIAL CLASS AND COMPONENTS OF S.J.

Ages	S.J. scale components	Mean scores		Total <i>W</i> score as percentage of total <i>M</i> score	N	P
		<i>W</i>	<i>M</i>			
10	A maximum score: 40	9.3	13.7	67.7	34	.008
14		17.6	33.4	52.6	26	
10	<i>aA</i> maximum score: 80 AE	23.2	35.2	64.7	34	.06
14		30.5	56.6	53.9	26	

\* These are shown only to indicate the scale of magnitude involved; the significance test was based on rank-ordering of matched pairs.

The top one is based exclusively on S.J. set *A*, which was common to both age-levels; the lower one includes the two different scales as a whole, i.e., *aA* and *AE*. The two modes of reckoning yield results that are compatible though somewhat different in magnitude.\* As indicated roughly in the percentage column (which refers to gross totals), the *W* children's vocabulary at age 10 was on the average two-thirds of that of their *M* peers, declining to just over one-half by the age of 14. The *P* computed from set *A* scores is highly significant, whilst the other one borders on the conventional level (two-tailed tests).

In evaluating this result, it must be kept in mind that the *W* boys were deliberately selected from among those restricted to a working class milieu, whereby the effect of such an environment (comprising both home and school) is likely to be

\* A detailed analysis of matched group data suggests that the easier scales at each age-level discriminate more sharply. Now, as *A* is the more difficult scale at 10, and the easier one at 14, it follows that *A* exaggerates the contrast, while *a* and *E* in combination underrate it; the true value is probably intermediate. In any case, the differences obtained are always partly a function of the particular word list employed.



specially accentuated. As against this, it should be pointed out that in this case the *best* (on a non-verbal test) of a relatively low-status group of *W* boys were compared with the *weakest* among a broader range of *M* boys, and this works in the opposite direction, i.e., it is likely to attenuate the discrepancy observed.

In general, the findings, expressed in the most cautious terms, clearly support the view that working class boys remaining within a working class environment tend to lose further ground in vocabulary performance between the ages of 10 and 14.

The writer is greatly indebted to Dr. A. W. Heim and Miss K. P. Watts for making the Self-Judging Vocubular Scale available for this investigation.

## REFERENCES.

- ANASTASI, A. (1960). *Differential Psychology*, 3rd ed. New York: Macmillan.  
BERNSTEIN, B. (1961). *Social Class and Linguistic Development: A Theory of Social Learning* in A. H. Halsey, Jean Floud and C. A. Anderson (eds.): *Education, Economy and Society*. New York: The Free Press.  
EELS, KENNETH, DAVIS, ALLISON, HAVIGHURST, ROBERT J., HERRICK, VIRGIL E., and TYLER, RALPH (1951). *Intelligence and Cultural Differences*. Chicago: Univ. of Chicago Press.  
RAVENETTE, A. T., and KAHN, J. H. (1962). Intellectual ability of disturbed children in a working-class area. *Brit. J. Soc. Clin. Psychol.*, 1, 208-212.  
MACARTHUR, R. S., and ELLEY, W. B. (1963). The reduction of socio-economic bias in intelligence testing. *Brit. J. Educ. Psychol.*, 33, 107-119.  
HEIM, A. W., and WATTS, K. P. (1961). A preliminary study of the self-judging vocabulary scale. *Brit. J. Psychol.*, 52, 175-186.

# SUMMARIES OF RESEARCHES REPORTED IN DEGREE THESES

## The Social Adjustment of Most Accepted and Least Accepted Children in Junior Schools.

By D. J. BELFIELD

(Abstract of Thesis submitted for the Degree of M.Ed., University of Manchester, 1963)

The main aim of the research was to investigate the relationship between social acceptance and social adjustment in children's school groups.

The method was to compare the behaviour and attitudes of seventy 'accepted' and seventy 'non-accepted' boys and girls. These children were selected by sociometric test from twelve classes, containing 479 children, drawn from six primary schools situated in a small urban community. Ages ranged from 9 to 11. The classes were unstreamed and confined almost entirely to children of the same age range. The sociometric test required three positive choices in each of three situations and one negative choice. A child's social acceptance score was the sum of the unweighted choice scores less the sum of unweighted negative choices he received. Limits, above and below which the children were selected, were approximately the 84th and 16th percentile ranks.

The following measures were taken to assess social adjustment: (1) Bristol Social-Adjustment Guides were completed by class teachers (Stott, 1958). The Guides were scored as Stott suggested; indicators of maladjustment were given 2 points and those of unsettledness and miscellaneous nervous symptoms were given 1 point. Social adjustment scores were related to behaviour in the following arbitrary adjustment groups: 0-9 'normal' (N); 10-19 'disturbed' (D); 20+ 'maladjusted' (Mal.). 'Maladjusted' children who had dominant A, HA, XC, HC or U, W groupings were classified as 'aggressive' (A) or 'withdrawing' (W). Where there was no obvious tendency to either pattern, 'maladjusted' behaviour was classified as 'mixed' (M) (compare Lunzer, 1960).

(2) Positive and negative attitudes expressed towards, or felt as coming from, the environment were investigated by a sentence completion test of twenty-nine items. The responses were coded in accordance with the system outlined by Bene (1957).

(3) Manifest anxiety was measured by the Short Form of the Children's Manifest Anxiety Scale (Levy, 1958).

(4) Moray House I.Q. taken at 9+ years were available in the schools.

(5) Ability at games was assessed on the basis of teacher's ratings.

All tests were administered under normal school conditions by the class teachers concerned. Reliabilities were obtained from re-test after eight weeks. A similar reliability test was made in respect of teachers' ratings...

Social adjustment data obtained from the Bristol Social-Adjustment Guide was checked by teachers' ratings of (1) Frequency of misbehaviour; (2) Extent and severity of problem behaviour; (3) Adjustment; (4) Co-operation; (5) Interest in school work.

The score ratings were also used to examine the teachers' attitudes to 'aggressive' and 'withdrawing' behaviour. In addition, a short questionnaire was completed by the children and intended mainly as an additional check on the reliability of several of the attitudes expressed more freely in the sentence completion test.

A follow-up experiment was carried out a year later in junior and secondary modern schools to investigate changes in the sociometric status of most accepted and least accepted children, and to relate these, to changes in social adjustment.

The following techniques were used to compare data based on the various sub-groups.



The standard error of the difference between the mean scores of the two main groups, each  $N=70$ , was obtained from the large sample formula; using CR as a measure of significance. The small sample formula, with 't' as a measure of significance, was used when either group numbered less than 70. Frequencies were analysed by  $\chi^2$ . Overall differences involving more than two groups were tested by simple analysis of variance. Where the F ratio was significant, 't' tests were made between pairs of means.

RESULTS OF MAIN EXPERIMENT (1961).

(1)

TABLE 1

RELIABILITIES OF CHILDREN'S TESTS.

Test	Number retested	Observed reliability	Type of coefficient
Sociometric .....	39 39	$\rho = .81$ $\rho = .87$	Rank-difference correlation
Sentence completion	75	$C = .66$	Contingency. 3x3 table. Max. value .816.
Manifest Anxiety ..	73	$r = .64$ $SE = .068$	Product moment correlation
Questionnaire .....	73	$r^t = .87$	Estimated from $r_{\cos \pi}$

(2) Reliabilities of teachers' ratings varied from  $r = .82$  to  $r = .90$ . Correlations between Bristol Guide social adjustment scores and the teachers' rating of adjustment were: most accepted children:  $r = .69$ ; least accepted children:  $r = .75$ .

(3) The most accepted children tended to have good, and the least accepted children tended to have poor, social adjustment scores. The difference in mean score for the total sample was highly significant. Differences in mean social adjustment score within class groups were all in favour of the most accepted children. In nine classes they reached significance at the .05 level or higher.

(4)

TABLE 2

NUMBERS IN MOST ACCEPTED AND LEAST ACCEPTED SOCIAL ADJUSTMENT GROUPS.

Sociometric status group	Social adjustment group				
	Mal. A	Mal. W	Mal. M	D	N
Most accepted .....	3	0 4	1	12	54
Least accepted .....	22	11 37	4	16	17

(5) The teachers, while recognizing 'withdrawing' behaviour as abnormal, thought that 'aggressive' behaviour was a more serious problem in school.

(6) High and low sociometric status was related to factors other than social adjustment. Social assets, associated with most accepted children, were : attractive physical appearance, good health, above average I.Q., above average ability at games, and positive attitudes, particularly to other children and to school. The reverse of these assets was associated with least accepted children.

(7) The least accepted children, particularly the socially 'maladjusted,' were associated with a 'vicious circle' of negative attitudes.

(8) No significant sex difference was found in either sociometric status or social adjustment.

#### RESULTS OF FOLLOW-UP EXPERIMENT (1962).

(9) The sociometric status of the Majority of the most accepted and least accepted children was fairly constant.

(10) There were no significant differences in sociometric status and social adjustment between least accepted children followed-up in secondary modern schools and those remaining in junior schools.

(11)

TABLE 3

CHANGES IN SOCIAL ADJUSTMENT GROUP OF LEAST ACCEPTED CHILDREN.

1962 Social adjustment groups	1961 Social adjustment groups				
	Mal. A N=22	Mal. W N=11	Mal. M N=4	D N=16	N N=17
Mal. A N=10 .....	6	0	0	3	1
Mal. W N=5 .....	0	4	0	1	0
Mal. M N=3 .....	2	0	1	0	0
D N=14 .....	7	1	0	3	3
N N=33 .....	7	6	3	6	11
Not followed up N=5	0	0	0	3	2

(12)

TABLE 4

CHANGED SOCIOMETRIC STATUS OF LEAST ACCEPTED CHILDREN IN RELATION TO CHANGED SOCIAL ADJUSTMENT.

1962 Sociometric status percentile rank	1962 Social adjustment groups					
	Mal. A	Mal. W	Mal. M	D	N <sub>m</sub>	N <sub>dn</sub>
Mean .....	18.77	28.17	11.10	25.42	26.86	44.44
N	10	5	3	14	16	17



The changed (1962) sociometric status of the  $N_{dn}$  (previously 'disturbed' or 'normal') group differed significantly, at the .05 level, from that of the Mal. A, Mal. W, Mal. M combined, and D groups, but not from that of the  $N_m$  (previously 'maladjusted') group.

#### CONCLUSIONS.

The Short Form of the Children's Manifest Anxiety Scale was considered to have unsatisfactory retest reliability.

The results of the main experiment, particularly the association of high sociometric status with good, and low sociometric status with poor, social adjustment, support the findings of other research in which the samples of children investigated were comparable.

Northway's theory, that a child's acceptance in a social group is related to the degree and direction of his outgoing energy, was confirmed.

The association of a 'vicious circle' of negative attitudes with socially 'maladjusted' children supports the general theory that maladjusted behaviour is reinforced by a 'vicious circle' of negative behaviour.

In the follow-up experiment, the improvement in social adjustment of many previously 'maladjusted' least accepted children occurred under normal conditions of primary and secondary school organization and was apparently spontaneous. It is thought that this result was partly due to changes in classroom environment and to differences between teachers in severity of scoring the Bristol Guide.

The spontaneous improvement in social adjustment of the previously 'maladjusted,' least accepted children was not at the same time, associated with much improvement in their sociometric status. That this may follow, provided improved behaviour is maintained, is suggested by the almost average sociometric status of the  $N_{dn}$  group.

Further research is needed to confirm the findings of the follow-up experiment and also to investigate ways of helping children who remain maladjusted and least accepted in school.

### **The Relationship between Social Status, Interests and Educational Achievement Among Children in a Primary School.**

By HAROLD N. BRUNT

*(Abstract of Thesis presented for the Degree of Master in Education in the University of Manchester, 1963)*

#### TEST POPULATION AND METHODS.

The enquiry was pursued with a mainly sociological emphasis on a sample of 100 children (forty-nine boys and fifty-one girls) in a Church of England voluntary primary school.

The school was situated within a large industrial city, about two-and-a-half miles from the centre, in a district which was mainly residential, but had some light industry, good shopping facilities and busy trunk roads. Housing was very mixed, as were the occupations of the parents.

Social status scores were obtained by using a social status schedule constructed by Chapman. It was based on a numerical assessment of the main living room of the home.

Educational achievement was measured in two ways. The children were given an achievement score in the basic subjects (B.S.) by adding the scores obtained on two standardised objective tests, one in English and one in Arithmetic. Another achievement score was given for educational achievement in the informational subjects (I.S.), History, Geography, Science and Scripture. Marks in these subjects were scaled to objective test scores with a mean of 100 and standard deviation of 15



Each child completed a diary form (marked in quarter hours), for each day of the week to measure his interests. The diaries were completed over a period of a month to minimise boredom. The diaries were supplemented by questionnaires on the children's interests.

#### ANALYSIS OF DATA.

The interests recorded in the diaries were summarised and classed as either 'individual' or 'group'; individual interests were those performed alone or with parents, group interests those performed in the company of other children. Each of these categories was further subdivided into 'common' or 'deviant' interests. Those interests which, on a common-deviant continuum fell beyond the highest quartile of deviance for boys and for girls separately were classified as deviant, the rest as common. This gave four types of interests for boys and girls: Individual Deviant, Individual Common, Group Deviant and Group Common.

Correlation coefficients were calculated between the intelligence scores, the social status scores and the two achievement scores for all the sample, for sex and for age. Chi-squares were calculated to test the significance of the relationships between those children with high and low 'Chapman' social status scores on the one hand and those children manifesting the two types of interests classified as 'none,' 'deviant exclusively,' 'common exclusively' and 'common and deviant' for both the individual and group categories. Chi-squares were also calculated for high and low achievement (both B.S. and I.S.) against the same interest groups.

#### RESULTS.

The sample used had a 'Chapman' social status mean score of 66.6 which was slightly less than the median 67.5 for a normal distribution. The boys' mean score was 64.5 that of the girls, 69.1. Thus, the two sexes were not equated populations in terms of social status.

From the diaries the proportion of the weekly time spent was (for all the children): in bed 46 per cent., on personal matters 11 per cent., school 15 per cent., chores 2 per cent., individual interests 4 per cent., and group interests 22 per cent. Of the twenty-seven different interests mentioned, the boys noted twenty-one and the girls twenty-six.

Television and street games dominated the group interests of both sexes. Reading comics and books, pet keeping and indoor games were highly popular individual interests.

Positive correlation coefficients were found between social status, and (a) intelligence, (b) achievement (B.S.) and (c) achievement (I.S.), the highest being +.605 between social status and achievement (I.S.) for the girls, and the lowest being +.381 between social status and achievement (B.S.) for the boys.

The difference of the sex roles at this stage appeared to determine the importance of the effect, in the basic subjects, of social status in favour of the girls (+.584 as against +.381). The correlation coefficients for social status and the informational subjects were higher for both sexes (girls +.615; boys +.569), than those for the basic subjects. The girls again appeared to be relatively more affected than the boys. It should be remembered that the girls' mean social status was itself higher than the boys'. In both the B.S. and I.S. scores, the correlations with social status for the first year children were higher than those for the fourth year children (B.S. 1st year +.535, 4th year +.462; I.S. 1st year +.585, 4th year +.505).

Coefficients of correlation between social status and intelligence on the one hand and between social status and achievement (B.S.) on the other were +.541 and +.544; almost the same. No statistically significant sex differences were found in any of the coefficients. To nullify the effects of intelligence in the achievement scores partial correlation coefficients were calculated between the social status scores and the two achievement scores (I.Qs. held constant in both cases). These were +.249 ± .067 (B.S.) and +.264 ± .066 (I.S.), ( $P < .01$  in both cases), both comparatively high. This apparently improbable result may be due to the discrepancy between the sexes in social status and the anomalies in the population itself.



Significant chi-squares were found between the individual interests, both 'deviant' and 'common,' and social status; and between achievement (B.S.) and the individual 'deviant' interests.

#### INTERPRETATION OF RESULTS.

From the correlation coefficients it could be seen that environment plays a significant part in the primary school child's performance and progress in school.

From the chi-square data there appeared to be a distinct relationship between children of low social status and a lack of individuality. The relatively few individual interests of children of low social status mainly concerned playing alone, both indoors and out, and were usually associated with rejection by their peers. The lack of significance in the group interest categories is probably due to the preponderance of group activities among children in the primary school and to the changing pattern of these groupings.

#### ADDITIONAL OBSERVATIONS.

From the diaries it appeared that children with high achievement scores in both categories were those who had a combination of both deviant and common individual interests. They also engaged in the group deviant interests such as Sunday School and Church attendance, swimming, going to the park and visiting with parents, watching sport, collecting things, keeping pets and roller skating. Club membership was a group deviant interest applying only to the girls.

The preponderance of common activities with correspondingly more deviance in the first year children as compared with the fourth year children could be seen as a reflection of the socialization process as outlined by Talcott Parsons. The younger child is nearer the adult pattern and has absorbed in lesser degree the institutionalised "differentiation of status on non-biological grounds," created by the school. There is less stability in the peer groupings; the child is more anxious to comply with teacher attitudes, the latter being more easily identified with the parental pattern. This pattern was confirmed by the results obtained in the achievement I.S. and individual interest categories; more deviance is shown in the high social status groups. The traditional and more formal school organization and the teaching of academic skills inhibits 'intellectual' interests. The family of high social status, however, encourages just these in the family situation and the differences are sharpened in the formal school situation as between child and child at the younger age. Group interests so dominate the activities of primary school children that only in a much larger sample could one expect statistically significant differences to be shown.

### The Post-School Social Adaptation of Educationally Subnormal Boys.

By G. C. MATTHEW

(Abstract of Thesis submitted for the degree of Master in Education, University of Manchester, September, 1963.)

#### (1) AIM OF THE RESEARCH.

The aim of the research was to study the post-school social adaptation of sixty-two educationally subnormal boys, ex-pupils of a special school in a county borough. An assessment was made of the degree to which they succeeded in employment, conformed to the standards of behaviour prevailing in the society in which they lived, mixed freely with boys of their own age and showed competence in managing the skills of everyday life. The standards against which the boys were measured were provided by a matched sample of boys of average intelligence, ex-pupils of secondary modern schools in the same borough.

## (2) METHOD.

Information was obtained from school records, a questionnaire, employers' reports, the Youth Employment Bureau, the Employment Exchange and the Child Guidance Clinic.

In addition, the following research tools were constructed :

(a) *A Social Competence Test*. This test was partly standardised and provided a visual summing up of the competence in various social skills of each subject of the E.S.N. sample.

(b) *A Questionnaire* for use with both samples, seeking information on employment histories, range of friendships, leisure activities and the use made of the amenities of the town in which they lived. Replies were presented in contingency tables, and chi-square tests were used to test the significance of observed differences between the samples.

(c) *A Questionnaire* for use when interviewing employers of the E.S.N. sample.

(d) *A Scale of Success/Failure in employment*. This was based on the proportion of time subjects had been unemployed since leaving school and on present employers' reports. Results were expressed on a five-point scale. The E.S.N. sample was judged on the full scale, the secondary modern sample on a modified scale.

(e) *An Index of Stability of Character*. A preliminary experiment was carried out to ascertain which character attributes, in the opinion of employers and teachers, contributed most to a stable character. For the five attributes ranking highest in their pooled judgment, emotional maturity, honesty, industry, perseverance and reliability, rating scales were devised and standardised. The index of stability of character was based on the combination of these rating scales and results were expressed on a five-point scale. The E.S.N. sample was rated on the scale by a team of three raters.

## (4) RESULTS.

(i) *The two samples.*

	E.S.N. Sample	Sec. Mod. Sample
No. of Boys .....	62	62
Mean age on 1st Jan., 1962 ..	19 yrs. 11 ms.	19 yrs. 11 ms.
Age range .....	16 yrs. 9 ms.—23 yrs.	16 yrs. 11 ms.—23 yrs.
Mean I.Q. ....	70.39 (S.D.=10.18)	99.16 (S.D.=6.09)
I.Q. Range .....	50—96	90—110
Mean No. of Years out of School .....	4 years.	4 yrs. 7 ms.

(ii) *Comparison of family backgrounds*. There was no significant difference in the size of family or in the type of house lived in by the two samples. Incidence of broken homes was, however, significantly greater among the E.S.N. sample. The majority of the fathers of both samples were engaged on skilled, partly-skilled or unskilled work, a significantly greater proportion of the fathers of the secondary modern sample being engaged on skilled work.

(iii) *Employment history*. The overall picture of the employment history of the two samples showed some 60 per cent. of the E.S.N. sample and 97 per cent. of the secondary modern sample, giving satisfaction in their occupations. Some 20 per cent. of the E.S.N. sample had less stable employment histories, were subject to more frequent job changes and to more periods of unemployment. The remaining 20 per cent. of the E.S.N. sample and 3 per cent. of the secondary modern sample had very unsatisfactory employment histories, being unemployed for more than 50 per cent. of the time since leaving school.



In earnings, the E.S.N. sample, when employed, were shown to be capable of earning a standard wage. The average earnings were slightly lower than those of the secondary modern sample, but this was to be expected from the nature of the employment of the two samples. The majority of the secondary modern sample were engaged on skilled work and the majority of the E.S.N. sample on unskilled work. The lower occupational status of the E.S.N. sample did not influence job satisfaction or relations with employers. The majority of the E.S.N. sample were happy in their work and did not find work discipline oppressive. There was evidence that leg-pulling was a common experience of both samples, but more of the E.S.N. sample found it a source of anxiety at the time. The E.S.N. sample tended to seek jobs nearer home. There was, however, no significant difference between the samples in the number of boys who showed complete independence of family by taking a job which necessitated living away from home.

Some 13 per cent. of the E.S.N. sample were working for sympathetic employers who accepted their shortcomings. Motives were not, however, entirely altruistic. Some were working for their fathers and some for sub-standard wages.

The E.S.N. sample had a greater burden of adverse experience to contend with. 38.7 per cent. came from broken homes; 16.13 per cent. came from homes in which unemployment and dependence on public support were viewed with equanimity.

(iv) *Factors contributing to success in employment.* Measured intelligence within the limits of the I.Q. range of the E.S.N. sample and level of attainment in the basic subjects were not associated significantly with success/failure in employment of the nature undertaken by the E.S.N. sample.

Stability of character, as measured by the index, was associated significantly with success in employment.

(v) *Leisure Activities.* The E.S.N. sample proved to be the less gregarious sample, to have fewer friends and to prefer passive viewing to active participation in recreational activities. A proportion, however, did conform to the pattern set by their intellectually superior fellows and their limited intelligence did not appear to handicap them unduly in the organisation of their leisure time.

The investigation revealed the existence of many cases of gross failure to conform to the normal pattern—boys who were completely friendless, who showed no ability or will to organise their leisure time and who were content to pass much of their spare time as passive viewers at the cinema or before the T.V. set.

(vi) *Social Competence.* The test of social competence and the section of the questionnaire dealing with budgeting of earnings revealed among the E.S.N. sample many areas of incompetence in understanding the mechanics of everyday life.

(vii) *Delinquency.* 25.8 per cent. of the E.S.N. sample and 16.12 per cent. of the secondary modern sample had appeared at least once before the juvenile court. The difference was not significant.

#### (5) CONCLUSIONS.

The E.S.N. subjects who made adequate social adjustment merged inconspicuously into the workaday world; the partial failures and the failures were potential social problems. It must be questioned whether the existing social services are adequate to deal with the post-school problem of the second category and, if not, whether the size and urgency of the problem justifies the creation of a specialised service to cater for their needs during the transition period between school leaving and adulthood.

## BOOK REVIEWS

ARGYLE, MICHAEL (1964). *Psychology and Social Problems*. Methuen, pp. 232, 25s. (U.K.).

The author has divided his book into four Parts: Social Behaviour, Social Problems, Processes of Social Control, and the Study and Control of Social Change. In Part 2, he deals with Aggression, Delinquency and Crime, Mental Health, Racial and International Attitudes and Human Problems in Industry; in Part 3 with Child-rearing effects, Adult Socialisation, Selection, Leadership, Mass Communications and Behaviour in Social Organisations.

The book is logically planned, and written in a clear style. The argument is sophisticated and well supported by research evidence. A notable feature of this evidence is that it is well chosen and up-to-date. The central theme is that our growing knowledge of social psychology and of related disciplines is already enabling governments to have a greater and more precise influence on social change. With the probable increase of research in this field, this tendency to plan social change for the good of the community, will grow. This reviewer, while agreeing that this process is good, is a little uneasy about such power getting into evil hands, as in the time of the Nazis. Yet one does not ban carving knives for fear they might be misused! The book concludes with a discussion on the use of the methods of the social scientists for the examination of Utopian experiments and the planning of future ventures.

The various parts of even social psychology—as treated in this book—are rapidly becoming specialisations in themselves. It is, therefore, steadily becoming more difficult for writers to cover all the aspects with equal competency, and it is very much to the credit of the author that there is not a weak chapter in the book. There are, however, one or two minor points of criticism. In his examination of the twin studies on the effect of heredity compared with the effect of environment, the author does not make it clear whether identical twins (which are, of course, of the same sex) are being compared with fraternal twins who might be of different sexes. Again, might not the fact that 85 per cent. of the identical co-twins of juvenile delinquents are also delinquents, but only 75 per cent. for fraternal twins, be due to the closer relationship between identical twins rather than due directly to heredity? Or would the author reply that this influence is entirely hereditary in nature?

There are one or two slight inconsistencies in argument. On page 61 we are told that "if the twin study method is reliable, this suggests that inheritance plays little part in juvenile delinquency," but on page 66 we read: "given that delinquents come from different kinds of genetic backgrounds," and on page 62: "It is more likely that both a muscular physique and a general failure to form inner restraints (and therefore produce delinquents) are jointly inherited, or the result of common biological processes." Nonetheless, this is a book well worth buying.

R. R. DALE.

BANDURA, A., and WALTERS, R. H. (1963). *Social Learning and Personality Development*. New York: Holt, Rinehart and Winston, Inc., pp. ix+329, \$7.00.

"In this book we have outlined a set of social-learning principles that emphasise the role of social variables to a greater extent than existing learning theories and consequently appear more capable of accounting for the development and modification of human behavior." This is the authors' claim, manifestly justified by this book. The five chapters are: The Socio-Behavioristic Approach; The Role of Imitation; Reinforcement Patterns and Social Behavior; The Development of Self-Control; The Modification of Behavior. These fields are wide, but the coverage is exceptionally thorough—astonishingly so when the slimness of the volume is taken into account.



The authors are critical of the wholesale application of operant conditioning explanations to the whole gamut of human learning, especially where the acquisition of new responses, as against the occurrence and modification of previously learned acts, is concerned. The discussion of learning theory applied to human beings, is handled in sophisticated fashion and the shortcomings of direct reliance on results derived from lower animals (all those concerned with the education of young humans should applaud this) tellingly brought out. Here, much is made of the concept of 'modeling,' the experimental work of the authors—most of it previously published in journals, but very useful under one cover—providing a definite line of considerable explanatory power. Motives antecedent to such modelling are not clear to this reviewer and I am reminded of McDougall's use of 'imitation,' 'primitive sympathy' and 'suggestion.' The authors' own discussion of imitation and of observational learning is sophisticated and very useful and throughout the book they never lose sight of the role of conditioning in human activity, though other types of learning are stressed. A very fair approach.

The approach is essentially a comparative one, animal and cross-cultural evidence being cited throughout. The latter use of social anthropology is relatively naïve though refreshing in that the text, packed with experimental data and references (the bibliography of just under 600 references is very useful indeed), is not very easy to read at first sitting.

Later dynamic 'mechanisms' such as *Identification* and *Regression* are re-examined, together with an extended new look at some issues in the fields of behavioural deviance and change, culminating with some considerations of various methods of psychotherapy, viewed against the theoretical position of the authors.

To sum up: I found the book very stimulating. It is probably too sophisticated for the first year psychology student but should, I think, be required reading for later years and, indeed, all psychologists with interests in the three fields of learning theory, social psychology, and developmental psychology.

One minor correction: in the Index, under Behavioral modification appears "see also Psychotherapy." There is, in fact, no entry under "Psycho therapy," but the relevant information can be found under "Treatment methods." S. G. LEE.

BENNETT, J. (1964). *Rationality*. London: Routledge and Kegan Paul, pp. viii+123, 14s.

This monograph is the most recent in a series of "studies in philosophical psychology" under the editorship of R. F. Holland. Earlier volumes include: R. S. Peters' *The Concept of Motivation*; Alasdair MacIntyre's *The Unconscious*; and Norman Malcolm's *Dreaming*. The writers are all professional philosophers and are concerned, not with advancing the empirical findings of psychology, but with the linguistic and conceptual problems that may arise when we attempt to interpret or to explain or to theorise about these findings. Some authors such as Peters and MacIntyre, have devoted themselves to the analysis of relatively technical psychological concepts; others, like the writer of the present book, are more interested in the clarification of the non-technical psychological notions used in ordinary language by those who are not professional psychologists.

The author's aim, then, is to analyse—to explore the conceptual ramifications of—our notion of 'rationality.' In particular, he is concerned to tease out the philosophical implications of the common beliefs that there is a difference in kind between the intellectual capacities of humans and those of animals, and that this difference is largely the result of the fact that humans use language whereas animals do not. Mr. Bennett admits that this is "an extremely ambitious conceptual enquiry," but the very ingenuity of his method and the sheer *panache* of his style will impress even the most sceptical reader. We are invited to consider the behaviour of von Frisch's honey bees: they can communicate to one another the distance and direction from the hive of sources of sugar by performing different sorts of dance. The writer argues, however, that this form of communication, despite its superficial resemblance to human language, cannot plausibly be counted as either 'intelligent' or 'rational.' This leads him to develop an extended 'apian fable,' in which the



bees' behaviour is, hypothetically, made progressively more complicated until we are at last convinced of its rationality. Finally, in his concluding sections, he is highly critical of what he clearly regards as the conceptual ineptitude of some of our leading experimental psychologists: he is especially shocked by the linguistic muddles which, in his view, have resulted from use of the concept of 'insight' for the interpretation of certain sorts of animal problem-solving behaviour.

Many psychologists will be irritated by this book. Some will feel that Mr. Bennett might have learned more of modern developments in animal psychology before being so caustically critical; others that, even given a proper understanding of recent work, he has grossly over-emphasised the relevance to scientific psychology of linguistic and conceptual issues; and yet others that his obvious brilliance might better have been devoted to some less recondite topic. Whatever the case, here is a superbly well written and well argued book. Readers curious to know of the kind of contribution that might be made to psychology by the analytic philosopher should certainly consult it.

G. W. PILKINGTON.

BRITISH COUNCIL FOR REHABILITATION OF THE DISABLED (1963). *The Handicapped School Leaver*. British Council for Rehabilitation of the Disabled, pp. 171, 15s.

This Report is a full-scale consideration of the problems of handicapped children in England and in Scotland in their transition from school to work and life after school. As such it is disappointing. Besides inefficient secretarial handling (there is no date, the list of witnesses is incomplete, there is no account of the relationship of the Working Party to the Council or the Inner Working Party to the main Committee, or of the procedure in taking evidence) the Report is full of vague statements, and its recommendations are not well supported by evidence. The one piece of evidence printed in full, the special enquiry of the Ministry of Labour into the employment position in certain selected areas, is detailed, but gives no information about the 94 per cent. found to be satisfactorily placed in employment and very little about the rest. The satisfaction shown by this enquiry that everything is being done for placement by the Y.E.O.'s. or D.R.O.'s. reflects well on this service, but from the point of view of a Working Party enquiring into the welfare of individuals, it was not helpful. There are some uncalled for passages dealing with ascertainment for special education, and about special education itself, both in ordinary and special schools which do not reflect any special study of a matter which was outside the terms of reference of the Working Party.

When dealing with matters of major concern, the Conclusions are generally sound and unexciting. Everyone must have known them already. There is enough legislation: the matter is, why is it not used better? Educational guidance leading not only to placement in a job but to satisfactory life adjustment is a process, and must not be thought of as an action taking place before leaving school. Collaboration among professions is stressed, but the difficulty of carrying parents with the experts is not imputed to poor communications at younger ages. Half-baked ideas on vocational training in school are firmly rejected. There are good suggestions for conference action instead of advice being passed back and forwards. The Report sensibly prefers persuasion of parents to let their children stay at school after the latest age of compulsion rather than raising the school-leaving age for special schools when that for ordinary schools is raised by a year.

But no mention is made of the effect the C.S.E. may have on special schools. The usual suggestions are made for the development of welfare services to carry on the influences exerted, at great expense, by the education services after school leaving; but the devoted agitation for improvement of this serious weakness in our social provision is not given much evidence to use with M.P.s. or the public.

Possibly the dullness and weakness of the Report is due to its being the work of a Committee, most of whose members are in the business up to the neck, who can hardly criticise present practice very severely since it is in part of their own making. One can only wish it had done some field work to see how former pupils of special schools were living and what *they* thought their needs were.



CHUKOVSKY, K. (1963). *From Two to Five*. Translated and edited by Miriam Morton. Berkeley and Los Angeles: University of California Press, pp. xx+168, \$4.50.

Observations of young children's speech collected by the author are made the basis of interpretative comment of a general kind. The book is chiefly valuable for one thing—its exploration of the importance of nonsense verse to young children in reinforcing their newly gained knowledge of the common-sense order of things. For the rest, it makes rather dull reading since it puts forward as solemn opinions matters now very familiar and provides little new evidence and not much in the way of explanatory hypothesis. When we have seen Vygotsky and Luria elaborating and testing hypotheses in this field, it is uninspiring to follow these lengthy declarations of opinion even when in detail the comments are as pithy and as charming as these sometimes are.

J. N. BRITTON.

COLE, ROGER (1964). *Comprehensive Schools in Action*. Oldbourne, pp. 223, 25s.

Discussions on comprehensive schools have engendered more heat than light. This book, by the deputy headmaster of a comprehensive school, grew out of an M.A. thesis at Sheffield University, and takes much of the heat out of the topic by the 'soft sell' technique. The theme is the anatomy and excellence of the Comprehensive school.

The opening chapters consists of a brisk gallop through educational history, dull stuff at the best of times, and well suited to the gallop, culminating in the Spens and Norwood reports and selection by examination, a tripartite system, an impossible 'separate but equal' slogan and a situation bristling with invidious comparisons. Such a potted history has little value, yet is an effective way of demonstrating that the system was not planned, but was merely 'a happening' in the sense of a notorious interlude at the last Edinburgh Festival. In an age which pays tribute to research and planning, albeit lip-service, it is chilling to read: "The grammar schools existed in sufficient numbers to provide for something approaching 20 per cent. of the population. It was assumed that about this proportion of the pupils would require an academic education of the grammar school type."

The major part of the book is descriptive of comprehensive schools, mainly in London. Their diversity is stressed, and the two-tier system considered acceptable. Desire for large numbers to ensure a good-sized sixth form sounds unconvincing, and doubts about dealing adequately with the academically talented are aired several times.

Much of the really practical discussion on buildings, staff, time-tabling and grouping of pupils into knowable groups, whether vertically or horizontally, will be of value to those employed in posts of responsibility in the large schools. The advantages of large units are enumerated, such as better equipment and libraries, wider choice and use of sites and better ancillary services.

The material on comparative education is scrappy. Scotland is given as much space as U.S.A. Original sources have not been used and the heart of the problem, that there are far more places available in grammar-type education, is not dealt with.

The author has aimed his book at too many people. Those most likely to be satisfied with it are teachers working in large comprehensive schools now, and concerned with day-to-day problems.

J. G. MORRIS.

GAGE, N. L. (Edited by, for American Educational Research Association) (1963). *Handbook of Research and Teaching*. Chicago: Rand McNally, pp. ix+1,218.

The budding educational research worker for whom this Handbook was intended will, on reading it, meet some chapters of almost histrionic evangelism and others succinctly reviewing available evidence. He might be forgiven for spending some time looking for the tripartite framework outlined in the *Preface* before



realising that this had soon been discarded as the *Contents* list four sections. A further misconception might have been avoided if the volume had been entitled *Handbook of American Research on Teaching*, as with greater or lesser forthrightness most of the chapters ignore the work done outside U.S.A. On closing the Handbook, he might also be forgiven for thinking what a sorry state research on teaching appears to be in and that perhaps he should look for another area.

He would recall the seeming lack of relevance to research on teaching of most of the chapters in Section I and will perhaps wonder if the problem is not already difficult enough without making paradigms about it. He might speculate whether social studies teachers really are the educational evangelists they were painted, and why the chapter on *Research on Teaching the Visual Arts* was devoted to the justification of Art Education. He might recall that the chapter on *Instruments and Media of Instruction* read like a telephone directory without being re-assured by his note that Lumsdaine was cited some 161 times in 88 pages. (Are research workers more important as names rather than for their work and ideas?) He will be relieved that Getzels and Jackson undermined the Minnesota Teacher Attitude Inventory, as he had been disturbed to read that good teachers on the basis of MTAI score had high scores for hysteria and psychopathic deviation. Not only that, they taught for the money and not for the inefficient teacher's reason of wanting to use his own knowledge and experience. On trying to wed the vagaries of subjective assessment of behaviour with the subsequent precise and exhaustive statistical analysis which occurred in *Measuring Classroom Behaviour by Systematic Observation*, he might recall a comment later in the handbook: "Statistical sophistication cannot make up for what might be termed theoretical naivete in the concepts." He could struggle to relate the content of Chapter 8, *Testing Cognitive Ability and Achievement*, with its title, and then decide that it had been a platform for the author's personal opinions on vital matters in education. He might regret the absence of discussion on motor development in the chapter on *Nursery Education* and the similar lack with regard to remedial techniques in the chapter on *Reading*. Finally, he would find disquietening the explanation by Getzels and Jackson why they had to discard yet another method of classifying the research in their area, "regrettably, the greater number of studies were so atheoretical that we were forced to abandon this plan."

It is difficult to believe that it would be better to substitute mental health criteria for the present morass of educational aims in assessing effective teaching. Such substitution might result in existing teachers seeing themselves as poorly trained mental therapists producing well-adjusted illiterates. It is disturbing to note the considerable disagreement between headmasters, colleges and administrative supervisors on the assessment of the same group of teachers, for this must surely reveal fundamental criteria differences.

Similar discord is revealed by Henderson's view of the 'sterile theories of educational psychology which have squeezed all logic from the theories of methodology,' but then one wonders whether this is really the fault of the psychologists if, as Wallen and Travers suggests, "there has been a tendency in education to pick and choose elements from the behavioural sciences that appear to fit with the philosophical traditions of teacher education." Surely, a naughty habit in research?

On the other hand, the concise, if rapid, development of educational statistics byatsuoka and Tiedman lends weight to their criticism of the lack of theoretical background of much of the research on teaching. The down-to-earth presentation of the chapter on experimental design by Campbell and Stanley was also refreshing, and, recalling some of the other chapters in the Handbook, one might smile ruefully at their suggestion that research workers should be prepared for barren results in many of their experimental studies. The chapter on teaching methods' research was intriguing in the light of its author's view that the personality of the teacher was the over-riding factor. If this is true, how many alleged studies of method had been bedevilled by this fact?

Wallen and Travers believe that much of the blame for the lack of theoretically based research could be laid at the door of the graduate schools of education in the universities and the apparent chaos in studies of teachers' behaviour and attitudes provides considerable support for Getzels and Jackson's plea for detailed study of the



personalities of teachers as a group before any further attempt is made to differentiate among them. All teachers would be advised to make a mental note of *Charters'* warning that social background is contemporary as well as historical, in that long service in a particular school will mould a teacher's attitudes, behaviour and teaching methods so that they might be quite unsuitable for a different school.

First and foremost among the general impressions is the considerable excess of opinion over evidence on most issues in teaching. Probably next is the frequently reiterated plea for accurate, thoroughly devised research to provide the fundamental theory or theories on which to base teaching methods. There is a general complaint against the lack of an agreed criterion of teacher effectiveness and without such a criterion what has been the value of the studies purporting to examine personality, social and academic differences between groups of alleged good and poor teachers? If the teachers' colleges train and assess on the basis of one set of criteria, headmasters and supervisors evaluate teachers on two others and children toss in a fourth estimate of the good teacher, these agencies, though obviously not antipathetic, are equally not working efficiently together.

Of the Handbook itself the reader might think that it was as well to have read it all but with the 1,200 odd pages behind him he might also have quietly wished that the editor had referred certain authors to their titles and wielded the blue pencil more vigorously in other chapters. Certainly, some contributions could have contained more fact and less preaching, but others were more than sufficiently thought-provoking to convince the would-be educational research worker that here, indeed, was a stimulating area in which to work—and in so doing, the Handbook is valuable.

K. B. START.

GOLDSTEIN, EDWARD (1964). *Selective Audio-Visual Instruction for Mentally Retarded Pupils*. Illinois: Charles C. Thomas, pp. xv+96, \$5.25.

This book has 90 pages, apart from a Preface and Acknowledgments. Of these 90 pages, 6 go to a discussion of learning through 'vicarious experiences' (i.e., films), 20 to a history of methods and attempts to educate the 'educable mentally retarded,' including the philosophy of 'occupational education,' all mainly from second-hand sources; 19 more cover the history of films and the use of films in teaching. The preparatory work for the research, the experimental design, the carrying out of the experiments, including a pilot study, the results, statistical analysis in detail, and the discussion and conclusions are all disposed of in the next 19 pages. Lest we miss any important detail, the remaining 24 pages contain the bibliography, questionnaires, objective test, the letter asking the co-operating teachers to fix times for the author to visit their schools, and information about the subjects, and their raw scores on the tests.

Three experimental groups (thirty-two subjects in all) were taught with a film as part of the lesson, a control group was taught the same material without the film. The same objective-type test on the material was given before and after the lesson. The author claims to show that the groups did not differ in I.Q. and age, but the analysis of variance in which he shows this cannot be checked, since he gives only the mean I.Q. and the range, and the mean and the range of ages. Checking would be necessary. For a Chi-square test (page 56) and a t-test (page 58) which could be checked, I could not arrive at his answers.

A complex analysis of variance, shown in full on page 58 is apparently based on Edwards' method (A. L. Edwards, 1950, *Experimental Design in Psychological Research*, pp. 295, *et seq*) and seems to be appropriate for his data. I have done this analysis, taking the number of groups as four (three experimental and one control), and as two (experimental and control), which is what the author seems to have done, but cannot arrive at anything like the same figures, except for the main totals. The allocation of degrees of freedom is very odd, also.

It is altogether a "horrid warning" of a book which as a review copy might have a certain limited usefulness for showing M.A. candidates what not to do. It would be over-finicky perhaps to mention the use of  $\chi^2$  for  $X^2$  and of E for  $\Sigma$ , numerous misprints and/or spelling errors, and the fact that the Appendix lettering in the text does not agree with that on the Appendices.

AGNES CRAWFORD.



HARRIS, DALE B. (1964). *Children's Drawings as Measures of Intellectual Maturity*. London: George G. Harrap and Co., Ltd., pp. 367, 70s.

The main emphasis of this book is indicated by its sub-title, "A revision and extension of the Goodenough Draw-a-Man Test" and its dedication to the memory of Florence L. Goodenough provides a further pointer to its general orientation. It is avowedly the work of a disciple rather than of a critic. The author has most ably met one of his aims, namely "to present a comprehensive survey of the literature on children's drawings in this country and abroad." Not only are over 500 articles and books reviewed, but these are listed in the bibliography in a way which makes it very easy to look up any reference in the text; and the name and subject index are done with similar competence. Since the presentation of past studies is, moreover, detailed enough to enable the researcher to make his own critical assessment, it is perhaps unreasonable to wish that the author had been more stringent in his evaluation of research designs, statistical methods and conclusions reached. It is probably more justified to regret that there are no suggestions for further research; for example, one would have thought there is a need for longitudinal studies of these revised and extended scales, all the more so in view of the lack of such studies with regard to Goodenough's original test.

These criticisms, however, do not seriously detract from the great value of the book to the student of children's drawings, for whom it is likely to become a major source book. It consists of two parts: the first and major part, reviews past studies, both historically and with respect to the use made of children's drawings for clinical and projective purposes. Then follows a description of the revised scales (there are three, drawing a man, a woman and the self) their reliability, validity and standardisation. In conformity with current procedures, the standard score method has been adopted (using a mean of 100 and standard deviation of 15). Tables are provided for converting raw scores into standard scores and the latter into percentile ranks.

The second part consists of the Test Manual. It contains instructions for administering and scoring the scales together with twenty-four sample drawings which make it possible to check one's own scoring with that of the author's. In the Appendix the scoring for a further forty-two drawings is given. The scales have been extended to cover a total of 73 scoring points (practically a 50 per cent. increase on the original test); this, together with the much more detailed description of how to judge each item, should make for greater reliability and for greater consistency between different scorers as well as markedly improving the ease of scoring.

Since the book provides both a comprehensive review of children's drawings and a fully illustrated and well-produced handbook for the revised and extended Draw-a-Man Test, its rather high cost seems justified. M. L. KELLMER PRINGLE.

KAHN, J. H., and NURSTEN, J. P. (1964). *Unwillingly to School*. Pergamon Press, pp. 187, 15s.

The increasing interest, in recent years, in the problems of children who refuse to go to school has resulted in a number of studies of 'school phobia' (or, as some prefer to call it, 'school refusal') from differing standpoints. These studies are so widely scattered in learned journals that a comprehensive book on the subject would not be untimely. This book, however, written jointly by a psychiatrist and a psychiatric social worker, although clearly written, reasonably free from jargon and well illustrated with case studies, is somewhat disappointing. The main theme is obscured at times by prolonged discussions of topics which would more appropriately be dealt with in such detail in a general work on child guidance. The references to the literature are rather disjointed, there being little attempt at a synthesis or a critical appraisal of previous studies. Although the authors stress the importance of a scientific approach to clinical problems, they frequently make sweeping generalisations and do not give us a systematic study of a group of cases.

In the first part of the book the writers distinguish between the problems of truancy and school phobia. The second part deals with the social, medical and clinical services available for the treatment of school phobia, and in the final section there is a discussion of methods of treatment within the child guidance clinic.



The need for a multi-disciplinary approach is stressed throughout, and there is little in the chapter on "The Role of the Psychologist" to which educational psychologists could take exception, though they may be surprised to learn (page 120) that "the psychologist's work in the school is based mainly on the cognitive level" and elsewhere (page 68) that "psychology is the study of normal mental processes, but academic psychology deals mainly with intellectual or cognitive aspects."

The authors, while recognizing that 'school phobia' is an umbrella term, state categorically that its cause lies within the child and within the family relationships, and advocate a psychoanalytic approach to clinical treatment. Educational factors receive scant attention, and little weight is given to relationships within the school—"alterations in the educational process are unlikely to have a positive effect." The book lacks a general index and there are a number of misprints.

Despite its shortcomings, the book is interesting to read, and the clinical insight of the authors will certainly help to give the lay reader, for whom the volume seems to be chiefly intended, a greater sympathy with the problems of emotionally disturbed children.

MAURICE CHAZAN.

KAPLAN, ABRAHAM (1964). *The Conduct of Inquiry*. San Francisco: Chandler Publishing Co., pp. 428, price:

This is a comprehensive, well-written and reasonably simple introduction into the philosophy and the logic of science. There are nine sections dealing, respectively, with methodology, concepts, laws, experiment, measurement, statistics, models, theories and explanation, and any psychologists wanting to find out painlessly what the middle of the road and sensible philosopher has to say about these topics could not do better than to go to this book. The style is readable and humorous, and clarity is one of the author's prime concerns. He is obviously well acquainted with psychology and the behavioural sciences generally and time and again his examples are taken from the psychological literature. Indeed, he makes a special point of discussing the application of principles of scientific methodology to the behavioural sciences in separate sub-sections after having dealt with their application to other sciences.

Would it be worthwhile for psychologists to read any book of this kind? There can be no general answer but I would be somewhat doubtful myself as to whether the budding scientist is greatly helped by this philosophical exercise. "Those who can, do; those who can't, teach" is a cruel saying but it does embody some degree of truth, at least in this field. Few of the outstanding psychologists that I have known have been particularly interested in, or benefited from, books such as this, and psychologists particularly knowledgeable in this field have usually had rather disappointing research careers. However, teachers who feel that their students should know something about the general principles of scientific enquiry could safely use this book; it may not turn the students into expert research workers, but it will at least not teach them the wrong things.

H. J. EYSENCK.

LEWIS, M. M. (1963). *Language, Thought and Personality in Infancy and Childhood*. London: Harrap, pp. 256, 18s.

This book, from the pen of Professor M. M. Lewis, will repay careful study. It is meticulously written, and the facts are well presented. Some of the material is already familiar to us in the author's earlier works, especially in *Infant Speech*. In this volume, however, the account of the development of speech and language has been brought up-to-date by the inclusion of research carried out by many workers in recent years, and linked up with general growth in infancy and childhood. Additional data which have accumulated from investigations into the effects of language impairment, as studied in individuals suffering from varying degrees of deafness, are also considered. Such results have yielded findings which are invaluable in themselves, but they also throw light on the development of speech and language in the hearing child.



Professor Lewis discusses development in Infancy, Early Childhood and Later Childhood. The whole gives a clear and comprehensive picture of the general characteristics of these periods, with a more detailed study of development of speech, language and personality at each stage. The section on Infancy is a penetrating account, which has been carefully assembled and which is well illustrated with many examples, of the very gradual changes which occur in speech and language. The author draws a useful distinction between cognitive and orrectic development, and this analysis of children's remarks forms one of the most interesting discussions in the book.

When impairment of language is present, retardation generally occurs. Lewis suggests that the retardation is not necessarily due only to poverty of language, but may arise from "the inevitable special orrectic relationship between a deaf child and his family." (page 75), a factor which may also be applicable to children in residence in institutions.

The work of Piaget and his followers is referred to frequently, and more especially in the chapter on Language and Concrete Thinking in the section on Later Childhood. This chapter is certainly worth special attention as it is a valuable assessment of some of Piaget's conclusions. The author discusses the view, also held by Piaget that the child's use of words may run ahead of his understanding and be a hindrance rather than a help in concrete thinking. Lewis expresses it thus: "Instead of thought through language, the child may become adept in manipulating verbal and numerical forms that replace thought and even hinder his development" (page 180). This theme is developed in a discussion of its practical implications for education. A section then follows on how language can promote concrete thinking, and eventually leads the author to 'internalised speech,' and the work of Luria, Vygotsky, Osgood and others.

As can be concluded from the examples quoted, the book is comprehensible and scholarly, and places in perspective much of the recent work in the field discussed. References are given for further reading in a separate list, and extra notes for each chapter are also appended.

The book captures the attention. It is clearly expressed and very readable. It can be recommended to the student of general psychology and education, as well as to those interested in the specific problems considered.

MARY COLLINS.

MASON, STELLA E. (1963). *Signs, Signals and Symbols*. London: Methuen, pp. xii+212, 35s.

This volume reprints a rather mixed bag of twenty-three papers which were presented at a conference held by the College of Speech Therapists in 1961. Outstanding are those by three guest speakers, Haas, Trim and Fry, who between them, justify in some detail the claim that the modern approach to linguistic studies can contribute a good deal to speech pathology and therapy. The essence of this approach is to see language not as "an indefinite agglomeration of individual facts," but as a hierarchical system of contrastive patternings, "a system consisting of several levels and organised in such a way that units which function on one level combine to form the units on the level next above." Thus, in his helpfully clear exposition of linguistic analysis on the (lowest) phonemic level, W. Haas shows how a large number of distinctive sounds can be grouped together, functionally, as allophonic variants of the same phoneme; it is, in fact, the phoneme ("a recurrent bundle of distinctive features") which serves to convey meaning within the sound system of any given language, and not the individual sound as such. Similarly, in his admirably lucid account of what is known about the encoding and decoding processes in speech, Professor D. B. Fry makes clear that it is relative and not absolute differences between acoustic features which are significant for the reception of speech. What rational justification can there be, then for treating then dyslalic child by "the patient teaching of individual sounds and the correct pronunciation of individual words"? It would seem that (as J. L. M. Trim puts it) "in many cases of infantile dyslalia, the child has constructed for himself a well-balanced phonological system



which is related to that of the adult world about him, but less complex"; and this hypothesis is explored in three highly interesting papers by R. E. Simms, L. M. Hartley and P. A. E. Grady, who describe and comment on their pre-diagnostic analysis of the language-system of an eight-year-old dyslalic boy. Nor, of course, is dyslalia the only speech disorder in which the linguistic approach may help towards more exact diagnosis and a proper assessment of priorities in treatment. In arguing the desirability of preliminary "analysis of the linguistic system which is being operated by the patient" in cases of developmental disorder, or of "determination of the linguistic levels affected, and to what extent" in cases of traumatic or degenerative disorder, Trim offers valuable and penetrating suggestions in regard to stammering, aphasia, and disorders of nasality.

A number of the other contributions seem to me too insubstantial to merit preservation in book-form, but it would be unfortunate if, as a result, the importance of these six papers were to be overlooked. They indicate a promising field for inter-disciplinary co-operation in which it is to be hoped that psychologists, psychiatrists, and medical specialists will play a full part, along with speech therapists and linguistic scientists.

FRANK WHITEHEAD.

ROTTER, JULIAN B. (1964). *Clinical Psychology*. London: Prentice-Hall International, pp. xv+112, 12s. 6d.

It is a little difficult to decide the audience for which this little paper-back has been written. In some places it is extremely elementary, in others a fair sophistication in psychology would be required for adequate understanding. It is said to be aimed at the 'introductory student.' Some of the problems discussed do not apply to practice over here: for example, whether the psychologist should see private patients of his own; and others are not live issues here any more: for example, whether clinical psychology should be an intuitive art or factual science. The book might be worth a quick reading by a student who wants to know something of the pre-occupations of clinical psychology, but the serious student would need to supplement it with other texts. The author has supplied lists of books which might be read to supplement each chapter.

RALPH HETHERINGTON.

STAATS and STAATS (1963). *Complex Human Behavior*. New York and London: Holt, Rinehart and Winston, pp. xiii+546, \$7.50.

This lengthy textbook has a strong Skinnerian flavour and is aimed at students of educational psychology. Its purpose is "to explore various experimental and naturalistic observations of complex human behaviour in terms of learning principles and thereby offer a relatively general conception of how the physical and social environments may shape human behaviour." The conceptual framework is expounded in the first half of the book. First, there is a useful statement of an operational approach to the study of behaviour. Then, some dozen 'learning principles' are listed, each tied to a simple prototype experiment. These principles hinge round the concepts of conditioning, reinforcement, generalization and discrimination, chaining, and drive. The framework is completed by a Skinnerian account of language behaviour and development. The second half of the book looks through this framework at a range of topics and cites many recent psychological studies. Successive chapters are concerned with personality, human motivation, social interaction, child development and training, experimental educational psychology, and behaviour problems and their treatment. The stress lies always on operational definitions and on those aspects of the topic which might be seen as exemplifying the key principles. Throughout, there is much that is worthwhile. The writing is lucid, and the book succeeds in showing that its selected principles have wide applications.

Of its kind, then, the book is good. However, this reviewer must confess a personal bias. He doubts the general utility of texts of this kind, especially for beginning students of psychology. The main ground for this doubt is that such texts deliberately focus more on their theory than on the behaviour which their



theory is designed to illumine ; and because of this, they frequently give the reader the impression of being led away from rather than into psychological enquiry. For example, this book raises the matter of children's ability to form word-endings appropriate to the word's part-of-speech. But it avoids questions which most students would ask about this interesting phenomenon, and seems concerned merely to argue that "grammatical word endings depend upon the formation of the appropriate response associations." And there the matter is left to rest. For this reviewer, this sort of treatment is so distinctly non-reinforcing that he would hesitate to recommend the book to any but the most sophisticated and persevering student.

In brief, the book does well what it sets out to do, namely, to delineate a consistent, fairly confident, and essentially Skinnerian approach to the broad study of human learning. But how useful it will be as a student text is a matter for doubt.

I. L. M. HUNTER.

STAFFORD-CLARK, DAVID (1964). *Psychiatry for Students*. George Allen and Unwin, Ltd., pp. 277, 35s.

This book is written primarily for post-graduate medical students and for general practitioners. It is written very much with practical clinical work in mind and little consideration is given to controversial issues. The tone of the book is human ; and again and again stress is made of the need for personal communication between psychiatrist and patient. The doctor is advised to study the patient first and the illness second, and to do this he is encouraged to see the mental condition as it exists subjectively for the patient. This may seem simple and obvious, but in practice the personal discipline of the clinician has to be quite deep for such communication to take place and the author is wise to reiterate this point.

The book is in two sections ; the major part, dealing with adult Psychiatry, is followed by a brief and somewhat condensed treatment of Child Psychiatry by Gerard Vaughan. The book concludes with an Appendix on Clinical Psychology by Jessie Williams. Dr. Stafford-Clark has covered the adult field with skill and fluency. A student will find a dependable and up-to-date summary, written without strong bias or prejudice, and the book will, therefore, serve as a good introduction as it is both lucid and fluent.

If the generic term student is intended to aim at post-graduate psychologists who are training as clinical or educational psychologists, the authors have written in too limited a manner, and one suspects that it is really a medical reader whom they have in mind.

The psychologist looks in vain for some discussion of concepts, of methods, and of research findings in areas where the psychologist has invaded the former medical province. Numerous examples of these omissions can be cited and mention of the following topics will serve to illustrate this point, and I hope, give guidance as to the value of the book to Psychologists. In the first place, the researches of the Maudsley Institute, however controversial they may be, are sufficient to constitute a challenge to psychiatric thought in some areas, and they merit discussion. But no mention is made of this work with regard to suggestibility, hysteria, or the personality dimensions of neuroticism and psychoticism. There has also been a failure to look at the work of the Neo-Behaviourists and the implications of this work for example, in conditioning treatment and Behaviour Therapy. In the Section on Child Psychiatry, Dr. Vaughan discusses the brain damage syndrome in its relationship to learning and behaviour problems and he has an extremely brief section on Educational Difficulties. Any psychologist would feel that inadequate attention has been paid to these difficulties having regard to the contribution they make to child maladjustment, and to their overlap with neurotic conditions. The author writes without awareness of the recent challenge made to the Concept of retardation by Professor Vernon.

Jessie Williams makes clear statements of the Psychologist's role both in the adult field and in that of Educational Psychology. She has had little space to do her task and has summarised the Psychologist's contribution to understanding in the areas of intelligence testing ; personality assessment, and brain damage, and she includes comments on recent research.

ENID M. JOHN.



VALETT, R. E. (1964). *The Practice of School Psychology*. New York: John Wiley, pp. 340, no price.

This book should interest educational psychologists in this country. At first glance it appears to be rather naïve and its style somewhat ponderous. Nevertheless, the author has something to say. The American school psychologist is not quite the counterpart of the British educational psychologist, just as the school systems do not correspond, but much in the book is apposite to the profession over here.

Part I is headed 'Theoretical Considerations.' A good deal of this deals with the impact of psychological theory on professional practice. This is an important topic, but its treatment here is inadequate and is really outside the scope of the author's main theme, which is the practical aspect of the psychologist's function.

The rest of the book is more down to earth and appears to be written out of a great deal of experience. Part II, dealing with 'Professional Relationships', makes this abundantly clear. Part III, 'Some Technical Problems' is the longest section and deals with problems of psychological evaluation, the assessment of educational achievement, psychological reports, the identification of and provision for individual differences and problems of counselling. This section is illustrated with a good deal of case material. Part IV, among other things, deals with problems in administering the service and future development of the profession.

In view of the relatively small number of educational psychologists in this country, a British book like this could hardly be produced with their specific conditions in mind. As the first in the field, this volume is worth recommending in spite of its American background and obvious limitations. N. E. WHILDE.

#### EDITORIAL NOTICE.

#### SELF-BINDING CASE FOR THIS JOURNAL

Several times recently, readers have asked whether it would be possible for us to produce a temporary binding case for journals. Messrs. Modern Bookbinders, Ltd., of Walpole Street, Blackburn, have now produced their 'Cordex' patent self-binding case to hold nine issues of the *Journal*, covering three years. A sample binder has been tried out for some time and found to be wholly satisfactory.

The binder is made of red book cloth, with the title in gold lettering on the spine. It is obtainable from the makers at 12s., including the price of postage. A remittance should be sent with the order.

## PUBLICATIONS RECEIVED

The mention of a book in this list neither implies, nor precludes, a later review.

- CALVIN W. TAYLOR (Edit.) (1964). *Creativity: Progress and Potential*. McGraw-Hill, pp. xiii+241, 54s.
- EDWARD GOLDSTEIN (1964). *Selective Audio-Visual Instruction for Mentally Retarded Pupils*. Springfield, Illinois: Charles C. Thomas, pp. xv+96, \$5.25.
- B. I. MURSTEIN (1964). *Theory and Research in Projective Techniques*. John Wiley and Sons, Ltd., pp. xiii+385, 64s.
- MARJORIE GRAHAM JANIS (1964). *A Two-year-old goes to Nursery School*. Tavistock Publications, pp. xi+156, 25s.
- SVEN MARKE and G. EBERHARD NYMAN (1963). *Perception of Parental Identification Parental Dominance, and Anxiety in Young Adults*. Lunds Universitets Arsskrift, pp. 47.
- ABRAHAM KAPLAN (1964). *The Conduct of Inquiry*. San Francisco: Chandler Publishing Co., pp. xix+428, no price given.
- BARBEL INHELDER and JEAN PIAGET (1964). *The Early Growth of Logic in the Child*. Routledge and Kegan Paul, pp. xxv+302, 40s.
- JOHN GABRIEL (1964). *Children Growing Up*. University of London Press, Ltd., pp. 367, 30s.
- H. S. WILLIAMS (1964). *Technical Education in Australia: Collected Papers*. Perth, Australia: Technical Publications Trust, pp. 88, 10s.
- NORMAN GEKOSKI (1964). *Psychological Testing*. Illinois: Charles C. Thomas, pp. xiii+300, no price given.
- HERBERT A. CARROLL (1964). *Mental Hygiene*. New Jersey: Prentice-Hall, Inc., pp. vii+408, no price given.
- HERBERT MARCUSE (1964). *One-Dimensional Man*. Routledge and Kegan Paul, Ltd., pp. xvii+230, 42s.
- W. DAVID WILLS (1964). *Homer Lane*. George Allen and Unwin, Ltd., pp. 275, 40s.
- COUNCIL FOR CULTURAL CO-OPERATION (1963). *Civics and European Education at the Primary and Secondary Level*. pp. 150, no price given.
- H. J. EYSENCK (Edit.) (1964). *Experiments in Motivation*. Pergamon Press, pp. viii+424, £5.
- CHESTER A. LAWSON and MARY ALICE BURMESTER (1964). *Programed Genetics*. Boston: D. C. Heath and Co., pp. vi+330, 32s.
- ROGER COLE (1964). *Comprehensive Schools in Action*. (Oldbourne Book Co., Ltd., pp. 223, 25s.
- JOSEPH M. WEPMAN and RALPH W. HEINE (Edits.) (1964). *Concepts of Personality*. Methuen and Co., Ltd., pp. xxix+514., 70s.
- L. P. BRADFORD, J. R. GIBB and K. D. BENNE (Edits.) (1964). *T-Group Theory and Laboratory Method*. London: John Wiley and Sons, Ltd., pp. xii+498, 75s.
- HARRY GUNTRIP (1964). *Healing the Sick Mind*. London: George Allen and Unwin, Ltd., pp. 223, 13s. 6d.
- JOHN A. DAHL, MARVIN LASER, ROBERT S. CATHCART and FRED H. MARCUS (Edits.) (1964). *Student, School, and Society*. San Francisco: Chandler Publishing Co., pp. xx+389, \$3.50.
- ERICH KAHLER (1964). *The Meaning of History*. New York: George Braziller, Inc., pp. viii+224, \$5.00.
- D. YOUNG (1964). *Non-Readers' Intelligence Test*. London: University of London Press, Ltd., pp. 27, 3s. 6d.
- M. L. J. ABERCROMBIE (1964). *Perceptual and Visuo-Motor Disorders in Cerebral Palsy*. The Spastics Society, Medical Education and Information Unit, in association with Heinemann (Medical Books), Ltd., pp. 136, 25s.



- JOHN P. DE CECCO (1964). *Educational Technology*. New York: Holt, Rinehart and Winston, pp. xiv+479, \$4.25.
- WILGA M. RIVERS (1964). *The Psychologist and the Foreign-Language Teacher*. University of Chicago Press, pp. viii+212, 30s.
- JULIAN B. ROTTER (1964). *Clinical Psychology*. London: Prentice-Hall International, pp. xv+112, 12s. 6d.
- PIERRE MALE (1964). *Psychétherapie de L'Adolescent*. Paris: Presses Universitaires de France, pp. 257, F.10.
- RICHARD R. SKEMP (1964). *Understanding Mathematics*. University of London Press, pp. 309, no price given.
- JOHN E. WATSON (1964). *Intermediate Schooling in New Zealand*. New Zealand Council for Educational Research, pp. xvi+460, 37s. 6d.
- MURRAY LEVINE and GEORGE SPIVACK (1964). *The Rorschach Index of Repressive Style*. Illinois: Charles C. Thomas, pp. xvi+164, \$7.50.
- JACK H. KAHN and JEAN P. NURSTEN (1964). *Unwillingly to School*. London: Pergamon Press, pp. xv+187, 15s.
- V. P. VARMA (1964). *Studies in the Philosophy of Education*. Agra: Lakshmi Narain Agarwal, pp. 338, Rs. 16.00.
- FLUGEL and WEST (1964). *A Hundred Years of Psychology*. London: Gerald Duckworth, pp. 394, 30s. Also in paperback by Methuen.
- L. S. HEARNshaw (1964). *A Short History of British Psychology*. London: Methuen, pp. xi+331, 35s.
- DAGOBERT D. RUNES (1964). *The Disinherited and the Law*. New York: Philosophical Library, Inc., pp. vii+79, \$3.00.
- DON C. CHARLES (1964). *Psychology of the Child in the Classroom*. New York: The Macmillan Co., pp. ix+86, 9s. 6d.
- BLAIR and JONES (1964). *Psychology of Adolescence for Teachers*. New York: The Macmillan Co., pp. vii+118, 11s. 6d.
- STRANG and MORRIS (1964). *Guidance in the Classroom*. New York: The Macmillan Co., pp. vii+118, 11s. 6d.
- NORMAN K. HENDERSON (1964). *Statistical Research Methods in Education and Psychology*. London: Oxford University Press, pp. vi+150, 30s.
- CHARLES J. BRAUNER (1964). *American Educational Theory*. London: Prentice-Hall International, pp. ix+341, 48s.
- ELWYN EDWARDS (1964). *Information Transmission*. London: Chapman and Hall, pp. 133, 15s.
- MARGARET MEAD (1964). *Continuities in Cultural Evolution*. London: Yale University Press, pp. xvi+471, 63s.
- EDWARD J. GREEN (1963). *The Learning Process and Programmed Instruction*. New York: Holt, Rinehart and Winston, pp. x+228, 36s.
- ERNST MARTIN (1964). *Grundformen des Gegenstandsbezuges im Unterricht*. Bern: Francke Verlag, pp. 410, S. Fr. 28.
- GEORGE A. MILLER (Edit.) (1964). *Mathematics and Psychology*. New York: John Wiley and Sons, pp. x+295, 42s.
- CLAUDE KOHLER (1963). *Les Déficiences Intellectuelles Chez L'Enfant*. Presses Universitaires de France, pp. 268, F. 12.
- WHITAKER and LIEBERMAN (1964). *Psychotherapy through the Group Process*. New York: Prentice-Hall, Inc., pp. x+305, 68s.
- FREDERIKSEN and GULLIKSEN (Edits.) (1964). *Contributions to Mathematical Psychology*. New York: Holt, Rinehart and Winston, pp. vi+189, no price given.
- THOMAS, CHESS, BIRCH, HERTZIG, and KORN (1964). *Behavioral Individuality in Early Childhood*. London: University of London Press, pp. xii+135, 21s.
- ARTHUR W. MELTON (Edit.) (1964). *Categories of Human Learning*. New York: Academic Press, Inc., pp. xvi+356, 68s.
- HEINZ HARTMANN (1964). *Essays on Ego Psychology*. London: The Hogarth Press and the Institute of Psycho-Analysis, pp. xv+492, 75s.
- D. BRUCE GARDNER (1964). *Development in Early Childhood*. London: Harper and Row, pp. x+358, 43s.

- DONALD ROSS GREEN (1964). *Educational Psychology*. London : Prentice-Hall, pp. viii+120, 12s. 6d.
- K. AUSTWICK (Edit.) (1964). *Teaching Machines and Programming*. Oxford : Pergamon Press, pp. vii+205, 42s.
- C. M. FLEMING (1964). *Teaching Psychological Analysis*. London : Methuen, pp. xiii+291, 12s. 6d. First published in 1958.
- IRVING SPERGEL (1964). *Racketville, Shuntown, Haulburg*. University of Chicago Press, pp. xxi+211, \$5.
- J. A. HADFIELD (1964). *Psychology and Morals*. London : Methuen, pp. ix+246, 9s. 6d. First published in 1923.
- H. J. EYSENCK (1964). *Crime and Personality*. London : Routledge and Kegan Paul, pp. xv+204, 25s.
- A. L. EDWARDS (1964). *Expected Values of Discrete Random Variables and Elementary Statistics*. London : John Wiley and Sons, pp. x+146, 45s.
- MARIA MONTESSORI (1964). *The Montessori Method*. Cambridge, Mass. : Robert Bentley, Inc., pp. xli+377, \$6.50.
- J. ROSWELL GALLAGHER and HERBERT I. HARRIS (1964). *Emotional Problems of Adolescents*. London : Oxford University Press, pp. x+210, 35s. First published in 1958.
- C. N. COFER and M. H. APPLEY (1964). *Motivation : Theory and Research*. London : John Wiley and Sons, pp. 958, 94s.
- NATHAN KOGAN and MICHAEL A. WALLCH. (1964). *Risk Taking*. London : Holt, Rinehart and Winston, pp. x+278, 64s.
- STEPHEN WISEMAN (1964). *Education and Environment*. Manchester : Manchester University Press, pp. xiv+202, 37s. 6d.
- NATIONAL ASSOCIATION FOR MENTAL HEALTH (1964). *Adolescence*. Report of the 20th Child Guidance Inter-Clinic Conference, April, 1964, pp. 40, 4s. 6d.





# INDEX TO VOLUME XXXIV, 1964

## Articles :

Adolescence, Role Conflict in : F. MUSGROVE .. .. .	34
Birth, date of, Backwardness and Educational Organisation : PHILLIP WILLIAMS ..	247
Career Prospects of a Group of Students in a College of Advanced Technology, the Social and Educational Background and Anticipated : RUTH RICE .. ..	256
Class Size and Co-ordinated Instruction : J. P. DE CECCO .. .. .	65
Composition, Written, at 10 years as an Aspect of Linguistic Development : O. C. SAMPSON .. .. .	143
Counselling, the Effect of on the Achievement of High Ability Pupils : G. SHOUK- SMITH and J. W. TAYLOR .. .. .	51
Hearing, the Attainments in English and Arithmetic of Secondary School Pupils with Impaired : D. C. WOLLMAN .. .. .	268
Imagery and Mental Practice : K. B. START and A. RICHARDSON .. .. .	280
Learning, Retention and Transfer of Concepts, Prompting and Feedback in the : M. C. WITTRICK and P. A. TWELKER .. .. .	10
Lie Scale, a, for the Junior Maudsley Personality Inventory : H. B. GIBSON ..	120
Maladjustment among Children in Schools for the Educationally Sub-normal : MAURICE CHAZAN .. .. .	292
Matrices, an Item Analysis of the Advanced : A. R. FORBES .. .. .	223
Motor Skill, Intelligence and the Improvement in after Mental Practice : K. B. START Performance in School Examinations, the Relationship Between Introversion- Extraversion, Neuroticism and : D. CHILD .. .. .	187
Personality of Pupils in a Secondary Modern School, a Comparison of Peer and Teacher Ratings : H. J. HALLWORTH and A. MORRISON .. .. .	285
Personality Ratings of Adolescents : A Study in a Comprehensive School : H. J. HALLWORTH .. .. .	171
Programmed Learning and Note-Taking, a Comparison at Two Age Levels : P. K. POPPLETON and K. AUSTWICK .. .. .	43
Readers, a Study of Some Cognitive Disabilities in Backward : K. LOVELL, D. SHAPTON and N. S. WARREN .. .. .	58
Readers, Backward, of Average Non-verbal Reasoning Scores, a Further Study of Some Cognitive and Other Disabilities in : K. LOVELL, E. A. GRAY and D. E. OLIVER .. .. .	275
Reversibility, Awareness of : its Effects upon Performance of Converse Arith- metical Operations : R. R. EIFERMANN and D. ETZION .. .. .	151
Science, the Factorial Nature of Attainment in Elementary : D. G. LEWIS .. ..	1
Scripture as a School Subject in Certain Co-educational Grammar Schools in Industrial South Wales, an Investigation into the Comparative Response of Boys and Girls : R. R. DALE and J. A. JONES .. .. .	132
Selection Procedure at Eleven, a Comparison of Rural Schools' Performance in a Test Controlled and a Test Free : H. G. ARMSTRONG .. .. .	237
Spelling, a Comparison of Two Approaches to the Teaching of : P. S. FREYBURG ..	178
Stereotypes in the Caretaking of English Children, a Study of : M. E. HEBRON and W. LONDON .. .. .	125
Teachers, Reasoning Ability and Personality Differences among Students : E. G. S. EVANS .. .. .	305
Teachers, Untrained and Trained Graduate, a Comparison of their Experiences during the Probationary Year : M. COLLINS .. .. .	75
Tests of Ability and Attainment, the Use of Group : N. FRANCE .. .. .	19
Valentine, Professor Emeritus, C. W., 1879-1964 : SIR CYRIL BURT .. .. .	219
Words, Adaptation to Level of Difficulty in Judging Familiarity of : A. W. HEIM, B. RAMSAY and K. P. WATTS .. .. .	109
Word Sounds by Young Children, the Analysis of : D. J. BRUCE .. .. .	158
Zero Error and Practice Effects in Moray House English Quotients : JOHN MACNAMARA	315

## Critical Comments :

Comments on the article by Magne and Parknas : M. D. VERNON .. .. .	204
---	-----

## Research Notes :

Personality and the Measurement of Intelligence : H. J. EYSENCK and P. O. WHITE	197
Social Class Differentials in Vocabulary Expansion : G. JAHODA .. .. .	321
A Statistical Re-analysis of Getzels' and Jackson's Data : R. W. MARSH .. ..	91
Puberty and Test Performance : A Further Report : J. D. NISBET, R. ILLSLEY, A. E. SUTHERLAND and M. J. DOUSE .. .. .	202

# Critical Notice :

The Education and Training of Teachers edited by G. Z. F. Bereday and J. A. Lauwerys :  
S. WISEMAN .. .. .

205

## Authors :

ARMSTRONG, H. G. : A Comparison of Rural Schools' Performance in a Test-Controlled and Test-Free Selection Procedure at Eleven .. .. .	237
BRUCE, D. J. : The Analysis of Word Sounds by Young Children .. .. .	158
BURT, SIR CYRIL : Professor Emeritus C. W. Valentine, 1879-1964 .. .. .	219
DE CECCO, J. P. : Class Size and Co-ordinated Instruction .. .. .	65
CHAZAN, MAURICE : The Incidence and Nature of Maladjustment Among Children in Schools for the Educationally Subnormal .. .. .	292
CHILD, D. : The Relationship between Introversion-Extraversion, Neuroticism and Performance in School Examinations .. .. .	187
COLLINS, M. : Untrained and Trained Graduate Teachers, a Comparison of their Experiences during the Probationary Year .. .. .	75
DALE, R. R., and JONES, J. A. : An Investigation into the Comparative Response of Boys and Girls to Scripture as a School Subject in Certain Co-educational Grammar Schools in Industrial South Wales .. .. .	132
EIFFERMANN, R. R., and ETZION, D. : Awareness of Reversibility : its Effects upon the Performance of Converse Arithmetical Operations .. .. .	151
EVANS, E. G. S. : Reasoning Ability and Personality Differences Among Student-Teachers .. .. .	305
FRANCE, N. : The Use of Group Tests of Ability and Attainment : A Follow-up Study from Primary to Secondary School .. .. .	19
FORBES, A. R. : An Item Analysis of the Advanced Matrices .. .. .	223
FREYBERG, P. S. : A comparison of Two Approaches to the Teaching of Spelling .. .. .	178
GIBSON, H. B. : A Lie Scale for the Junigr Maudsley Personality Inventory .. .. .	120
HALLWORTH, H. J. : Personality Ratings of Adolescents : A Study in a Comprehensive School .. .. .	171
HALLWORTH, H. J., and MORRISON, A. : A Comparison of Peer and Teacher Personality Ratings of Pupils in a Secondary Modern School .. .. .	285
HEBRON, M. E., and LONDON, W. : A Study of Stereotypes in the Caretaking of English Children .. .. .	125
HEIM, A. W., RAMSAY, B., and WATTS, K. P. : Adaptation to Level of Difficulty in Judging Familiarity of Words .. .. .	109
LEWIS, D. G. : The Factorial Nature of Attainment in Elementary Science .. .. .	1
LOVELL, K., GRAY, E. A., and OLIVER, D. E. : A Further Study of Some Cognitive and Other Disabilities in Backward Readers of Average Non-Verbal Reasoning Scores .. .. .	275
LOVELL, K., SHAPTON, D., and WARREN, N. S. : A Study of Some Cognitive and Other Disabilities in Backward Readers of Average Intelligence as assessed by a Non-Verbal Test .. .. .	58
MACNAMARA, JOHN : Zero Error and Practice Effects in Moray House English Quotients .. .. .	315
MUSGROVE, F. : Role Conflict in Adolescence .. .. .	34
POPPELTON, P. K., and AUSTWICK, K. : A Comparison of Programmed Learning and Note-Taking at Two Age Levels .. .. .	43
RICE, RUTH : The Social and Educational Background and Anticipated Career Prospects of a Group of Students in a College of Advanced Technology .. .. .	256
SAMPSON, O. C. : Written Composition at 10 years as an Aspect of Linguistic Development .. .. .	143
SHOUKSMITH, G., and TAYLOR, J. W. : The Effect of Counselling on the Achievement of High Ability Pupils .. .. .	51
START, K. B. : Intelligence and Improvement in a Gross Motor Skill after Mental Practice .. .. .	85
START, K. B., and RICHARDSON, A. : Imagery and Mental Practice .. .. .	280
WILLIAMS, PHILLIP : Date of Birth, Backwardness and Educational Organisation .. .. .	247
WITTRICK, M. C., and TWELKER, P. A. : Prompting and Feedback in the Learning, Retention and Transfer of Concepts .. .. .	10
WOLLMAN, D. C. : The Attainments in English and Arithmetic of Secondary School Pupils with Impaired Hearing .. .. .	268



## NOTES TO CONTRIBUTORS

1.—The Editor is glad to consider articles in the fields of educational and child psychology, whether written from the theoretical, clinical, experimental or psychometric standpoints. It is, however, desirable that statistical matter (formulae, tables, discussions of techniques, etc.) should be kept to a minimum, so that the *Journal* will be of interest to statistically untrained readers. Such concepts as the mean, percentiles, standard deviation, correlation, chi-squared, standard error, critical ratio, significance and reliability, and general or group factors, can be employed in the text. But more complex concepts and methods should be explained or confined to footnotes or an appendix, or published elsewhere.

Articles which have been, or are to be, published elsewhere cannot be accepted.

2.—*Length.* The usual range of length is from 2,500 to 4,500 words, and only very exceptionally will articles of more than 6,000 words be published. Short research reports of 500—1,000 words, such as outlines of theses, are welcomed. These should normally be submitted through the Head of the Department in which the research was carried out.

3.—*Arrangement.* Articles should be preceded by a summary or abstract of between 100 and 300 words, embodying the main conclusions. This should be typed directly after the title and author's name and place of work. A contents analysis, or list of headings, need not be included, but successive sections of the article should receive numbered headings, e.g.:

- I. INTRODUCTION.
- II. OUTLINE OF PREVIOUS RESEARCH.
- III. PLAN OF THE INVESTIGATION.
  - (i) *Subjects and materials.*
  - (ii) *Experimental Procedure.*
- IV. STATEMENT OF RESULTS.
- V. DISCUSSION AND CONCLUSIONS.
- VI. REFERENCES.

4.—*Typing.* All papers must be typed (double or  $1\frac{1}{2}$ -spaced) on one side of quarto or foolscap paper. There should be margins of at least one inch all round. Authors should consult articles in current numbers of the *Journal*, and copy their conventions in respect of capitalising the title and main section headings, italicising author's place of work and sub-section headings, indentation of paragraphs, the form of references and tables, etc.

5.—*Footnotes* should be few in number, and should be indicated in the text by using the signs \*, †, ‡, §.

6.—*References* should normally be indicated by date following the author's name in the text, e.g., Spearman (1904), or (Spearman, 1904, p. 160), or—where there are two or more publications in the same year—Burt (1939a). However, it is occasionally more convenient to number the references throughout, and to type numbers in the text above the line, e.g., Spearman.<sup>4</sup> When there are two or three references only, these may be placed in footnotes in the text, instead of in a separate section at the end.

The list of references at the end should always be in alphabetical order of author's names, and—for any one author—by order of date. The following standard form of references should be observed:

DEARBORN, W. F., and ROTHNEY, J. W. M. (1941). *Predicting the Child's Development*. Cambridge, Mass.: Sci.-Art.

(Note that Ltd., etc., after publisher's names, also the numbers of pages in the book, can be omitted.)  
BURT, C. (1939a). The relations of educational abilities. *Brit. J. Educ. Psychol.*, 9, 55-71.

(Note the absence of inverted commas and of capital initials in the titles of articles. The number of the volume is in arabic, not roman numerals. The part of the journal is not given, but page numbers are included in full.)

The World List of abbreviations (as given in the December issues of *Psychological Abstracts*) is used for journals.

7.—*Tables and figures* are costly to print, and should be used only when they save space and/or are essential for effective presentation. The same material should not be presented in a table and a graph. Each table or figure should be numbered (in arabic figures), on a separate sheet, and its approximate position indicated in the typescript. Drawings should be in Indian ink on heavy, unruled paper, the lettering on graphs being in pencil. Lines in tables should be ruled in pencil. Each table or graph should have a short title, sufficient for understanding without searching in the text.

8.—*Corrections.* The Editor cannot consider for publication articles which are seriously deficient in presentation. The typescript submitted (after preliminary consultation, if desired) should represent the final form in which the author wishes the paper to appear. Since changes in proof are costly, out of all proportion to the original setting, authors must be charged for any changes, insertions or deletions other than printer's errors).

9.—*Reprints.* Authors receive fifty copies of their papers free. Extra copies, in multiples of fifty, may be had at cost price if the order is given when the proof is returned.

**E. & S. LIVINGSTONE, LTD.**

**THE DEVELOPMENT OF THE INFANT  
AND YOUNG CHILD. Normal and  
Abnormal.**

*Second Edition.* By R. S. Illingworth,  
M.D., F.R.C.P., D.P.H., D.C.H.

370 pages. 126 illustrations. 30s.

"A splendid simplification and condensation of the original studies of Gesell and others adapted by the author for day-to-day clinical use and widely modified by his own large experience. The book is entertaining and it is practical; it is very well written; it should certainly be read by all who are engaged in the study of young children."

*Archives of Disease in Childhood.*

**CEREBRAL PALSY**

By A. Katharine Levin, M.S.A.O.T.,  
M.W.F.O.T., S.R.O.T.

132 pages. 104 illustrations. 30s.

The story of the pioneer years of occupational therapy in Scotland, and of the efforts made to alleviate serious disability among those handicapped by cerebral palsy.

**TEVIOT PLACE, EDINBURGH**

**THE  
BIRMINGHAM  
PRINTERS LTD**  
PRINTERS & BOOKBINDERS

*High Class  
Technical  
Publications*

**42-44 HILL STREET  
BIRMINGHAM 5**

**TELEPHONE : MIDLAND 0196-7**

**LEARNING**

*Winfred Hill*

'Teaching machines' and 'programmed learning' are phrases which are beginning to penetrate the consciousness of the man in the street. Even for students of psychology there are few simple descriptions of the theories that lie behind these modern techniques.

The present work by an American psychologist provides a sound and simple account of these theories.

University Paperback 9s. 6d.  
Hardbound 18s.

*Recent University Paperbacks*

**TEACHING**

*C. M. Fleming*

**A HUNDRED YEARS OF PSYCHOLOGY**

12s. 6d.

*J. C. Flugel*

21s.

**METHUEN**

& Co. Ltd. · 11 New Fetter Lane · London E.C.4